UNIVERSIDAD D CÓRDOBA

Facultad de ciencias

Departamento de Bioquímica y Biología molecular

Programa de Doctorado en Ingeniería Agraria, Alimentaria, Forestal y de Desarrollo Rural Sostenible

Identificación de las Familias de Factores de Transcripción WRKY y Proteínas VQ en Fresa y Caracterización del Papel de FaWRKY1 en la Respuesta de Defensa del Fruto frente a *Colletotrichum acutatum*

Identification of the WRKY Transcription Factor and VQ Protein Families in Strawberry and Characterization of The Role of FaWRKY1 in the Fruit Defense Response against *Colletotrichum acutatum*

Tesis Doctoral presentada por

José Garrido Gala

Directores

Dr. José Luis Caballero Repullo

Dr. Francisco Amil Ruiz

Córdoba, septiembre 2021

TITULO: *Identificación de las Familias de Factores de Transcripción WRKY y Proteínas VQ en Fresa y Caracterización del Papel de FaWRKY1 en la Respuesta de Defensa del Fruto frente a Colletotrichum acutatum*

AUTOR: *José Garrido Gala*

Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

https://www.uco.es/ucopress/index.php/es/
ucopress@uco.es

TÍTULO DE LA TESIS:

Identificación de las Familias de Factores de Transcripción WRKY y Proteínas VQ en Fresa y Caracterización del Papel de FaWRKY1 en la Respuesta de Defensa del Fruto frente a *Colletotrichum acutatum*.

DOCTORANDO/A: Ldo. JOSÉ GARRIDO GALA

INFORME RAZONADO DEL/DE LOS DIRECTOR/ES DE LA TESIS

(se hará mención a la evolución y desarrollo de la tesis, así como a trabajos y publicaciones derivados de la misma).

El licenciado José Garrido Gala ha desarrollado el trabajo de investigación titulado: "Identificación de las Familias de Factores de Transcripción WRKY y Proteínas VQ en Fresa y Caracterización del Papel de FaWRKY1 en la Respuesta de Defensa del Fruto frente a Colletotrichum acutatum" que constituye el tema de su Tesis Doctoral para optar al grado de Doctor.

Este trabajo de investigación se ha realizado en el seno del Grupo de investigación BIO-278, del Departamento de Bioquímica y Biología Molecular de esta Universidad de Córdoba, adscrito al ceiA3. Ha sido dirigido y supervisado por el Dr. José Luis Caballero Repullo -Catedrático de Universidad y miembro del citado Departamento de Bioquímica y Biología Molecular de la Universidad de Córdoba- y codirigido por el Dr. Francisco Amil Ruiz.

Durante el periodo de investigación en el que se ha desarrollado esta tesis, el doctorando ha adquirido una firme competencia investigadora, así como una amplia experiencia en diferentes técnicas de Biología Molecular y Bioinformáticas, además de demostrar una gran capacidad de aprendizaje y trabajo. Su formación le ha permitido optar a diferentes puestos de trabajo en distintos centros y empresas, que ha compaginado con la realización y finalización de su tesis: AGROMÉTODOS SA, como Técnico De Investigación Y Desarrollo; Instituto de Agricultura Sostenible (CSIC), como Titulado Superior De Actividades Técnicas y Profesionales; y, actualmente, investigador en Phytoplant Research SLU.

La investigación realizada se ha centrado, por una parte, en caracterizar genéticamente a las familias de los factores de transcripción WRKY y de las proteínas con motivo valina-glutamina (VQ) de la fresa diploide (*Fragaria vesca*) y octoploide (*Fragaria x ananassa*), obteniéndose resultados novedosos. Así, destacamos el descubrimiento de un nuevo tipo de proteínas VQ, presentes tanto en fresa como en otras especies vegetales y que hemos denominado "proteínas R-VQ". No menos importante, se ha estudiado el papel del gen *FaWRKY1* durante la infección del fruto de fresa frente con el hongo *Colletotrichum acutatum*, demostrándose su papel como regulador negativo en fresa, en contraste con su función como regulador positivo de defensa, previamente descrita en Arabidopsis frente a la bacteria patógena *Pseudomonas syringae*.

Durante el desarrollo de esta tesis doctoral las comunicaciones presentadas a congresos más destacables han sido:

Resistencia frente a *Colletotrichum acutatum* en frutos de fresa: prueba de función de genes mediante ensayos de expresión transitoria.
José Javier Higuera Sobrino, María Isabel Arjona Girona, Francisco Amil Ruiz, **José Garrido Gala**, Aymam Lekhbou, Juan Manuel Arjona López, Enriqueta Moyano, Juan Muñoz Blanco, Carlos José López Herrera, José Luis Caballero.
XVIII Congreso de la Sociedad Española de Fitopatología. Palencia. 2016. Oral.

El silenciamiento del gen *FaNPR3.1* de fresa (*Fragaria x ananassa*) genera resistencia a la infección por *Colletotrichum acutatum* y sugiere una función en defensa semejante a la de los genes *AtNPR3/AtNPR4* en Arabidopsis.
José J. Higuera-Sobrino, Francisco Amil-Ruiz, **José Garrido-Gala**, Ayman Lekhbou, Enriqueta Moyano, Isabel Arjona-Girona, Juan M. Arjona-López, José A. Mercado, Fernando Pliego-Alfaro, Juan Muñoz-Blanco, Carlos J. López-Herrera, José L. Caballero.
XXXIX Congreso de la Sociedad Española de Bioquímica y Biología Molecular (SEBBM). Salamanca. 2016.

Transient silencing of the *FaWRKY1* strawberry gene (*Fragaria x ananassa*) in fruit induces resistance to *Colletotrichum acutatum* infection. José J. Higuera-Sobrino, Francisco. J. Molina-Hidalgo, Isabel Arjona-Girona, Francisco Amil-Ruiz, **José Garrido-Gala**, Ayman Lekhbou, José A. Mercado, Fernando Pliego- Alfaro, Juan Muñoz-Blanco, Carlos J. López-Herrera, José L. Caballero.
Plant health sustaining mediterranean ecosystems Cordoba. 2017. 15th Congress of the Mediterranean Phytopathological Union. Póster.

The Strawberry FaWRKY1 Transcription Factor Negatively Regulates Resistance to *Colletotrichum acutatum* upon Fruit Infection.
Jose J. Higuera-Sobrino, **José Garrido-Gala**, Ayman Lekhbou, Francisco J. Molina-Hidalgo, Isabel Arjona-Girona, Francisco Amil-Ruiz, Jose A. Mercado, Fernando Pliego-Alfaro, Juan Muñoz-Blanco, Carlos J. Lopez-Herrera, Jose L. Caballero.
3rd International Conference “Plant Biotic Stresses & Resistance Mechanisms”. Viena. 2018. Póster.

Los artículos de investigación publicados han sido los siguientes:

Garrido-Gala, J., Higuera, J. J., Muñoz-Blanco, J., Amil-Ruiz, F., & Caballero, J. L. (2019). The VQ motif-containing proteins in the diploid and octoploid strawberry. Scientific Reports, 9(1), 4942. https://doi.org/10.1038/s41598-019-41210-4. (Impact Factor 3.998/Q1/Multidisciplinary)

Higuera JJ*, **Garrido-Gala J***, Lekhbou A, Arjona-Girona I, Amil-Ruiz F, Mercado JA, Pliego-Alfaro F, Muñoz-Blanco J, López-Herrera CJ and Caballero JL (2019) The Strawberry FaWRKY1 Transcription Factor Negatively Regulates Resistance to *Colletotrichum acutatum* in Fruit Upon Infection. Front. Plant Sci. 10:480. doi: 10.3389/fpls.2019.00480 *These authors have contributed equally to this work (Impact Factor 4.402/Q1/Plant Science)

Y los artículos de investigación que se encuentran actualmente en fase de escritura son los siguientes:

Garrido-Gala, J., Higuera, J. J., Muñoz-Blanco, J., Amil-Ruiz, F., & Caballero, J. L (2021). A comprehensive study of the WRKY transcription factor family in strawberry.

Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 22 de julio de 2021

Firma del/de los director/es

CABALLERO REPULLO JOSE LUIS - 30432037R

Digitally signed by CABALLERO REPULLO JOSE LUIS - 30432037R
DN: c=ES, serialNumber=IDCES-30432037R, givenName=JOSE LUIS, sn=CABALLERO REPULLO, cn=CABALLERO REPULLO JOSE LUIS - 30432037R
Date: 2021.07.22 10:51:12 +01'00'

Fdo.: José Luis Caballero Repullo

Firmado por AMIL RUIZ FRANCISCO - 30954974X el día 28/09/2021 con un certificado emitido por AC FNMT Usuarios

Fdo.: Francisco Amil Ruiz

Índice general

Resumen

La fresa cultivada es la fruta blanda de mayor importancia económica, con una producción mundial de más de 8,33 Mt, en la que España ocupa el sexto puesto en importancia, y ocupando una superficie de cultivo de unas 372.361 ha (FAOSTAT, 2018). Numerosas especies de hongos patógenos pueden producir enfermedades tanto de la planta como del fruto de fresa, siendo estas las de mayor importancia, tanto por su variedad como por su incidencia en distintas regiones productoras, siendo responsables de cuantiosas pérdidas económicas. La antracnosis es una de las principales enfermedades de la fresa cultivada, extendida globalmente y que afecta a todos los principales países productores. Está causada por hongos ascomicetos del género *Colletotrichum*.

En la presente tesis doctoral se ha abordado, por una parte, el estudio de dos familias génicas propias de plantas que codifican proteínas reguladoras de la transcripción: los factores de transcripción tipo WRKY y las denominadas "proteínas VQ", un tipo de proteínas cuyo papel regulador se ha identificado recientemente, capaces de establecer interacciones proteína-proteina con los primeros y modular su actividad transcripcional. Ambas participan en la regulación de diversos aspectos del crecimiento y desarrollo y, especialmente, en las respuestas de defensa frente a estreses bióticos y abióticos. Por tanto, se han identificado y caracterizado a los miembros de dichas familias génicas en la fresa, tanto en la especie modelo diploide, *Fragaria vesca*, como en el híbrido octoploide cultivado, *Fragaria* x *ananassa*, mediante un profundo y sistemático análisis en el que se usaron los datos genómicos actualizados de ambas especies y técnicas avanzadas de filogenómica y análisis de la expresión génica (RNA-seq y PCR cuantitativa). De este modo, se describe la evolución de ambas familias génicas y sus posibles funciones, destacando aquellas en relación a los mecanismos de defensa de la fresa frente a uno de sus principales patógenos, el hongo hemibiotrofo *Colletotrichum* spp.

Por otra parte, se ha profundizado en el papel del gen *FaWRKY1* en la respuesta de defensa de la fresa. En un trabajo previo realizado por nuestro grupo, se estudió el gen *FaWRKY1*, que codifica un factor de transcripción WRKY del grupo IIc y homólogo al gen *WRKY75* de *Arabidopsis thaliana*, revelando un papel en la respuesta de defensa de la planta frente a la antracnosis del fruto, causada por *Colletotrichum acutatum*. Mediante la técnica de agroinfiltración en fruto se consiguió tanto la sobre-expresión como el silenciamiento transitorios del gen *FaWRKY1*, hallándose que éste ejerce como regulador negativo de la resistencia del fruto de fresa frente a *C. acutatum*, en contraposición a su papel como regulador positivo de la resistencia frente al patógeno bacteriano *Pseudomonas syringae* pv. Tomato DC3000, determinado mediante complementación en *A. thaliana*. Se ha propuesto un modelo para explicar este comportamiento dual en los mecanismos de defensa frente a distintos patógenos, en el cual FaWRKY1 contribuiría a la regulación del crosstalk entre las dos principales rutas de defensa, mediadas por SA (ácido salicílico) y JA (ácido jasmónico), respectivamente.

Abstract

The cultivated strawberry is the most economically important soft fruit, with a worldwide production of more than 8.33 Mt, in which Spain ranks sixth in importance, and occupying a growing area of about 372,361 ha (FAOSTAT, 2018). Many species of pathogenic fungi can produce diseases of both the plant and the strawberry fruit, these being the most important, both for their variety and for their incidence in different producing regions, being responsible for large economic losses. Anthracnosis is one of the main diseases of the cultivated strawberry, spread globally and affecting all major producing countries. It is caused by ascomyte fungi of the genus *Colletotrichum*.

This doctoral thesis has addressed, on the one hand, the study of two classes of gene families typical of plants that encode transcription-regulating proteins: WRKY type transcription factors and so-called "VQ proteins", a recently identified type of regulatory proteins, capable of establishing protein-protein interactions with the former and modulating their transcriptional activity. Both are involved in regulating various aspects of growth and development and, in particular, in defense responses against biotic and abiotic stresses.

Therefore, the members of these gene families have been identified and characterized in the strawberry, both in the diploid model species, *Fragaria vesca*, and in the cultivated octoploid hybrid, *Fragaria x ananassa*, by a deep and systematic analysis in which the updated genomic data of both species and advanced techniques of phylogenomics and analysis of gene expression (RNA-seq and quantitative PCR) were used. This describes the evolution of both gene families and their possible functions, highlighting those in relation to the mechanisms of defense of the cutter against one of its main pathogens, the fungus hemibiotroph *Colletotrichum* spp.

On the other hand, it has delved into the role of the *FaWRKY1* gene in the strawberry defense response. In previous work carried out by our group, the *FaWRKY1* gene, which encodes a WRKY transcription factor of the IIc group, homologous to the *WRKY75* gene of *Arabidopsis thaliana*, was studied revealing a role in the plant's defense response against fruit anthracnose, caused by *Colletotrichum acutatum*. The fruit agroinfiltration technique achieved both transient over-expression and silencing of the *FaWRKY1* gene, finding that it acts as a negative regulator of strawberry fruit resistance against *C. acutatum*, as opposed to its role as a positive regulator of resistance against the bacterial pathogen *Pseudomonas syringae* pv. Tomato DC3000, determined by complementation in *A. thaliana*. A model has been proposed to explain this dual behavior in the mechanisms of defense against different pathogens, in which FaWRKY1 would contribute to the regulation of crosstalk between the two main defense pathways, mediated by SA (salicylic acid) and JA (jasmonic acid), respectively.

Agradecimientos

Algo tan complejo como una Tesis necesita de mucha dedicación propia, pero también implica tiempo, cooperación y el apoyo de muchas otras personas. Aunque suelo ser parco en palabras, según saben los que me conocen, quisiera dedicar unas pocas líneas para mostrar mi gratitud por la ayuda, mucha y variada, recibida en la realización de este trabajo.

Así, en primer lugar quisiera agradecer a mis directores, José Luis Caballero y Francisco Amil, por el esfuerzo y la dedicación que me han prestado a lo largo de estos años.

Por supuesto, a los miembros y compañeros que han pasado, o siguen, por el laboratorio, pero especialmente a Fran, Nico, Javi, José Javier, Laura Medina, Laura Diezma, Natali, Félix, Enri y Charo, con los que he compartido muchas horas de laboratorio y muchos momentos "Camera Café". ¡¡¡Menudos ratos hemos pasado!!! ¡Nos seguiremos viendo para contar batallitas!

También tengo que agradecer a todo el personal del Departamento de Biología Molecular y Bioquímica por su colaboración y apoyo durante el desarrollo de esta Tesis, particularmente al grupo de Emilio Fernández y Aurora Galván, así como al grupo de Jesús Jorrín, por tener las puertas abiertas cuando hacía falta.

Y sobre todo, agradezco y dedico esta Tesis a mi familia, ya que muchísimas de las horas invertidas en este trabajo han sido, casi siempre, a costa de vuestra compañía y atención, cosa a la que pondré remedio de inmediato.

¡Gracias a todos!

Capítulo 1.

Introducción general

Capítulo 1. Introducción General

1.1. Filogenia y evolución de *Fragaria* spp.

1.1.1. Rósidas

Las rósidas forman un clado bien diferenciado, constituido por más de 90.000 especies y 140 familias taxonómicas que suponen aproximadamente la cuarta parte de la diversidad del extenso grupo monofilético de las eudicotiledóneas (Wang et al., 2009;Soltis et al., 2019). Recientes estudios moleculares han conducido a que se las divida en dos grupos: Vitales (fam. Vitaceae, p.e., *Vitis*) y las eurósidas. Estos últimos, a su vez, se clasifican en fábidas (Fabidae, sin. eurrósidas I) y málvidas (Malvidae, sin. eurrósidas II). Las rósidas, junto con las *Saxifragales* forman el clado de las superrósidas (Group, 2016) (Figura 1). Estos términos sistemáticos responden a criterios cladísticos basados en estudios filogenéticos, más que los tradicionalmente taxonómicos, posibles gracias a la creciente disponibilidad de datos moleculares de diferentes especies.

Las rósidas presentan una amplia diversidad morfológica, bioquímica y ecológica que se traduce en gran heterogeneidad en cuanto a ciclos de vida y hábitats, por lo que encontramos especies herbáceas, leñosas, suculentas, acuáticas e incluso, parásitas. La razón de ésta diversidad puede encontrarse en una temprana radiación de los principales clades de rósidas, ocurrida en el Cretácico tardío (115-93 Ma), relativamente poco después de la aparición en el registro fósil, hace unos 125 Ma, de las primeras plantas con flor (angiospermas) y a continuación del evento gamma de triplicación completa del genoma ancestral de las eudicotiledóneas (117 Ma) que se cree contribuyó de forma decisiva a la evolución y diversificación del clado (Wang et al., 2009;Jiao et al., 2012;Chanderbali et al., 2017;Soltis et al., 2019) (Figura 2).

1.1.2. Familia Rosaceae

Dentro del orden Rosales encontramos a la familia Rosaceae, que incluye más de 100 géneros y alrededor de 3.000 especies (Hummer and Janick, 2009), entre las que se encuentran numerosos frutales de especial relevancia económica. Mientras que se han hallado fósiles de rosáceas datados en unos 90 Ma, se estima que la diversificación de la familia comenzó hace unos 100 Ma, con una rápida divergencia que originó a las subfamilias, seguida por una expansión gradual de las principales tribus que abarcó desde el Cretácico hasta el Paleoceno (Crepet et al., 2004;Xiang et al., 2017).

Las rosáceas se distribuyen en tres subfamilias, bien caracterizadas por su número cromosómico y tipos de fruto, que han sido redefinidas recientemente en base al

análisis filogenético: Amygdaloideae, Rosoideae y Dryadoideae (Xiang et al., 2017) (Figura 3). En la subfamilia Amygdaloideae, que incluye varias especies de interés agrómico con frutos principalmente tipo drupa o pomo, se han agrupado las subfamilias clásicas:

1. Amygdaloideae, con especies principalmente leñosas y frutos tipo drupa (ciruelo, cerezo, melocotonero, almendro). Poseen un número haploide de 8 cromosomas (n=8).
2. Spiraeaoideae, que incluye hierbas y arbustos de interés ornamental (Spiraea, Aruncus, Sorbaria y otras), con frutos tipo folículo (agregados). Poseen 9 cromosomas (n=9).
3. Maloideae, también abundante en especies leñosas, predominan los frutos tipo pomo (por ejemplo, manzano, peral, níspero). Poseen 17 cromosomas (n=17).

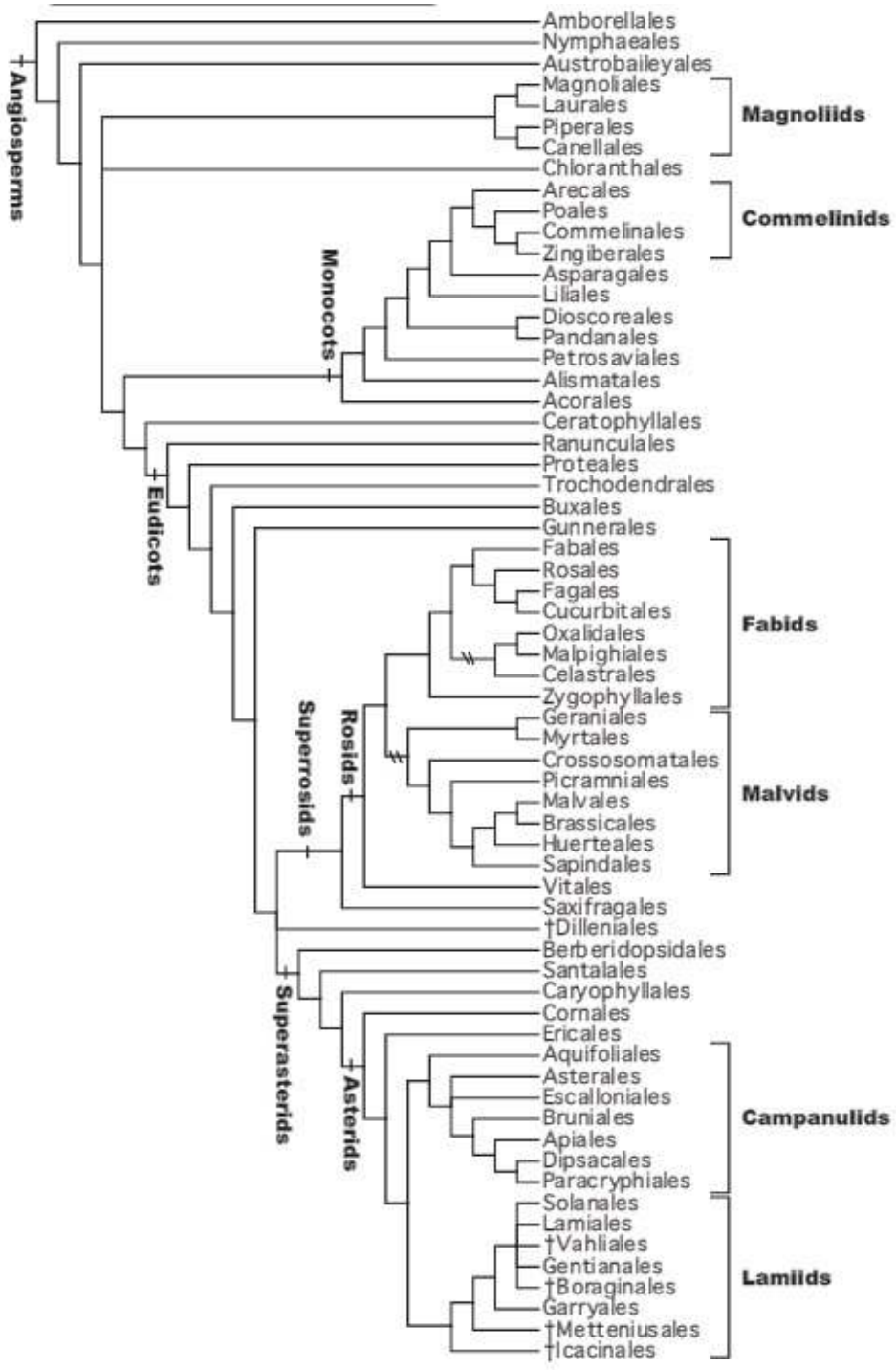
La subfamilia Dryadoideae (n=9) incluye a cuatro géneros de arbustos (*Dryas, Purshia, Cercocarpus* y *Chamaebatia*) caracterizados por formar frutos tipo aquenio o poliaquenio y por desarrollar nódulos fijadores de nitrógeno en sus raíces, mediante simbiosis con *Frankia*.

Por último, la subfamilia Rosoideae (n=7), que incluye a las rosas (género *Rosa*) y otros géneros de notable importancia económica, entre los que destacan aquellos con especies perennes y productoras de frutos agregados del tipo aquenio o drupa. Por ejemplo: *Rubus*, al que pertenecen varias especies de frambuesas, moras y zarzamoras; y *Fragaria*, con especies naturales e híbridas productoras de diversas variedades de fresas silvestres y cultivadas.

1.1.3. Género *Fragaria*

El género *Fragaria* (Fam. Rosaceae, subfam. Rosoideae, tribu Potentilleae, subtribu Fragariinae), incluye actualmente a 24 especies, descritas en base a su morfología, distribución geográfica y nivel de ploidía, principalmente (Folta and Davis, 2006;Liston et al., 2014;Sobczyk, 2018). Así, encontramos especies diploides (2n=2x=14), tetraploides (2n=4x=28), hexaploides (2n=6x=42), octoploides (2n=8x=56), decaploides (2n=10x=70) y tres especies híbridas naturales: *F. × bifera* Duchesne (diploides y triploides, 2n=3x=21), *F. × bringhurstii* Staudt (pentaploides, 2n=5x=35, hexaploides, 2n=6x=42 y nanoploides, 2n=9x=63) y *F. × ananassa ssp. cuneifolia* Staudt (octoploide, 2n=8x=56; híbrido natural de *F. chiloensis* y *F. virginiana*). Filogenéticamente, las especies de *Fragaria* se agrupan en dos clados principales: el clado “vesca”, con especies endémicas del continente americano y Eurasia, y el clado “asiático” o “chino”, formado por especies originarias de Asia. Las relaciones filogenéticas con algunas de las especies diploides, distribuidas por Asia y Europa, no han sido totalmente resueltas debido a la pobre resolución obtenida usando marcadores genéticos clásicos (genes nucleares y plastomas), lo que sugiere una divergencia evolutiva limitada dentro del género *Fragaria* (Njuguna et al., 2013;Sobczyk, 2018). Esto podría explicarse en base a su reciente aparición (en

términos geológicos) hace unos 2 Ma, mientras que se estima en poco más de 1 Ma la aparición de las especies octoploides (Njuguna et al., 2013) (Tabla 1 y Figura 4).



© 2016 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2016, **181**, 1–20

Figura 1. Filogenia de las Angiospermas, APG IV.

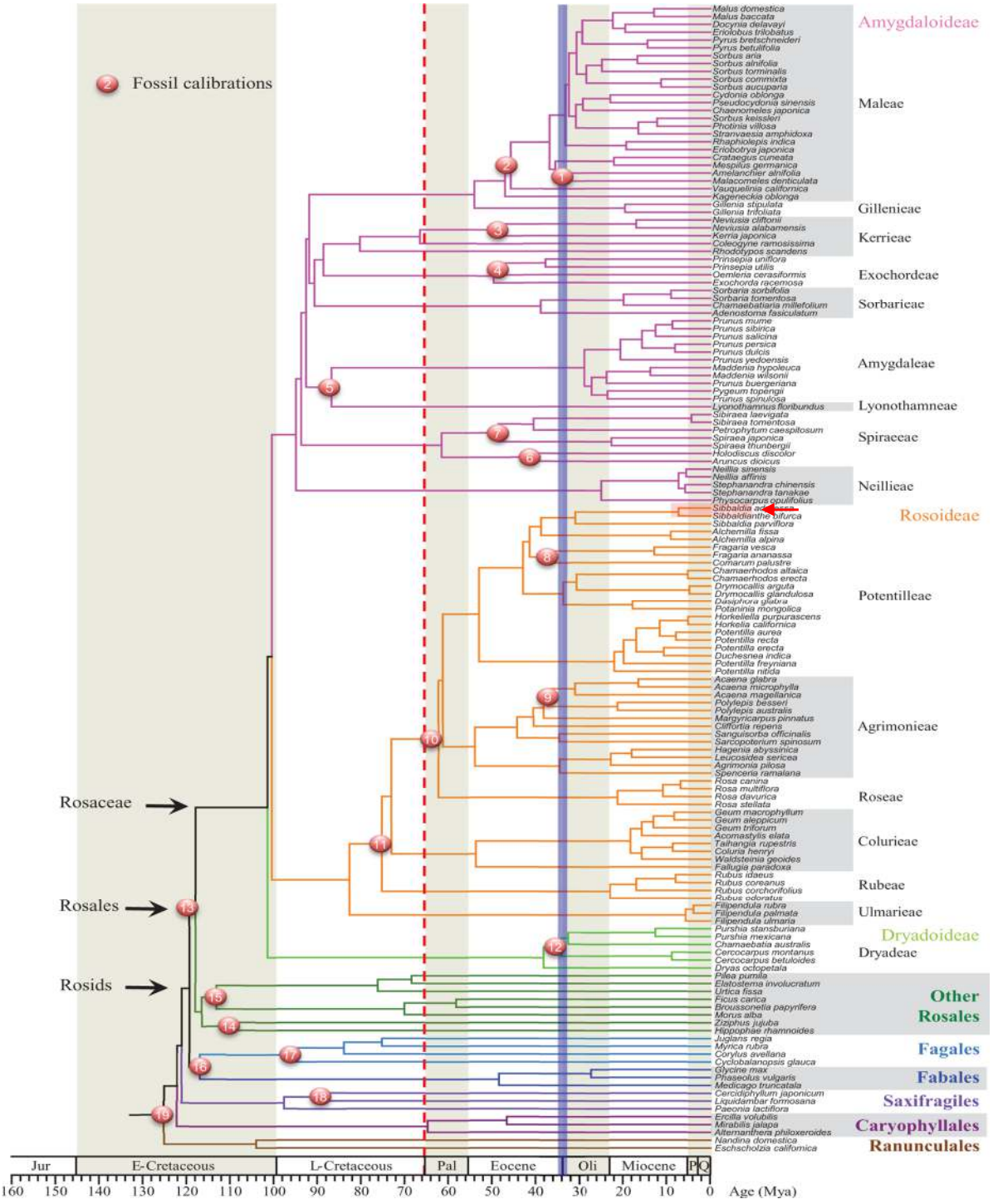


Figura 2. Cronograma y relaciones filogenéticas de las rósidas, mostrando la diversificación del clado, así como la filogenia de la familia Rosaceae en detalle. Una flecha indica la posición de *F. vesca* y *F. × ananassa* dentro de la tribu Potentilleae. Los globos numerados corresponden a puntos de calibración basados en el registro fósil. Figura adaptada de Xiang *et al.*, 2017.

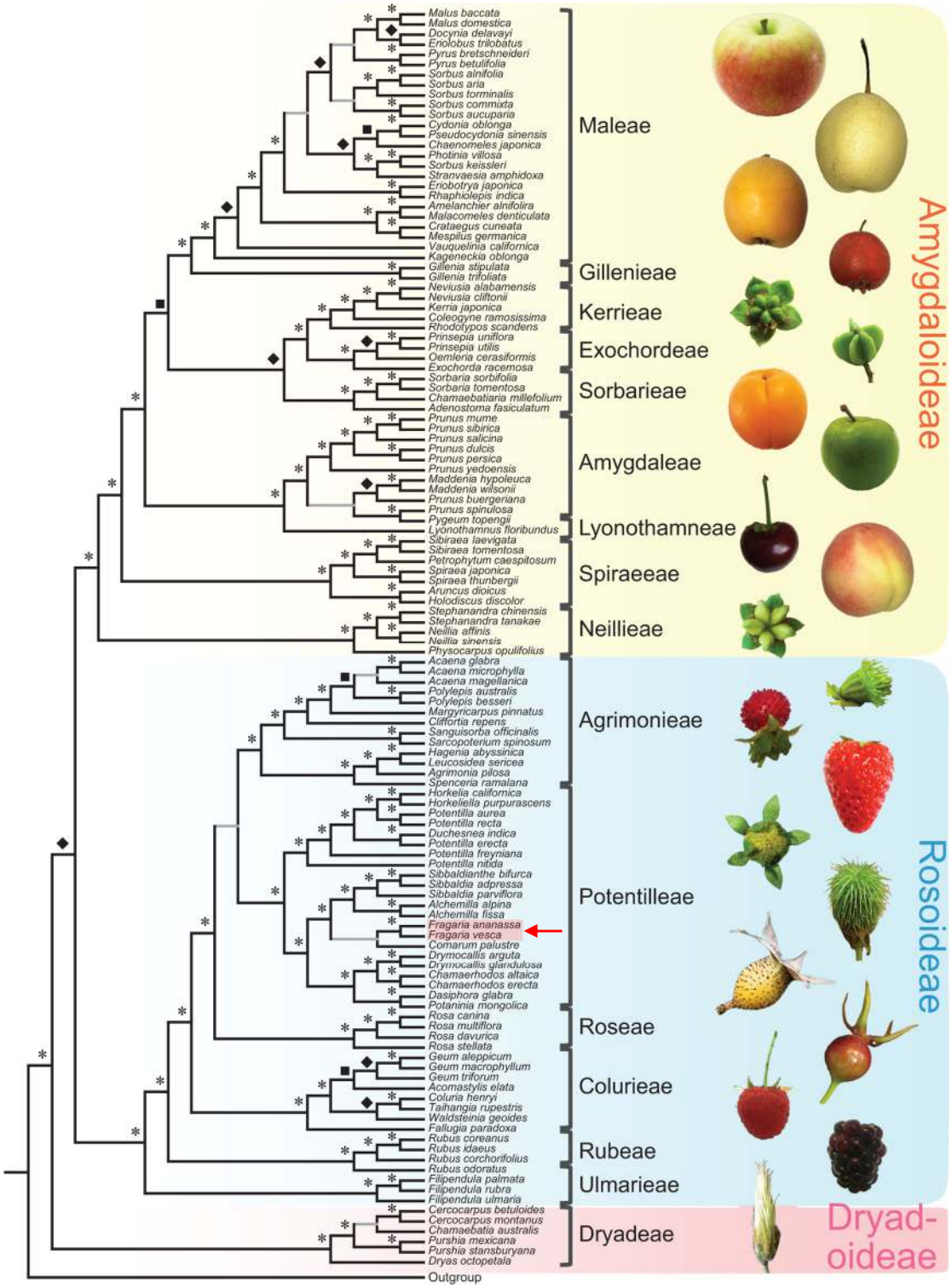


Figura 3. Filogenia de la familia Rosaceae en relación a los diferentes tipos de frutos encontrados dentro de las principales subfamilias y tribus, descritos en el texto. Una flecha indica la posición de *F. vesca* y *F. × ananassa* dentro de la tribu Potentilleae. Figura adaptada de Xiang *et al*., 2017.

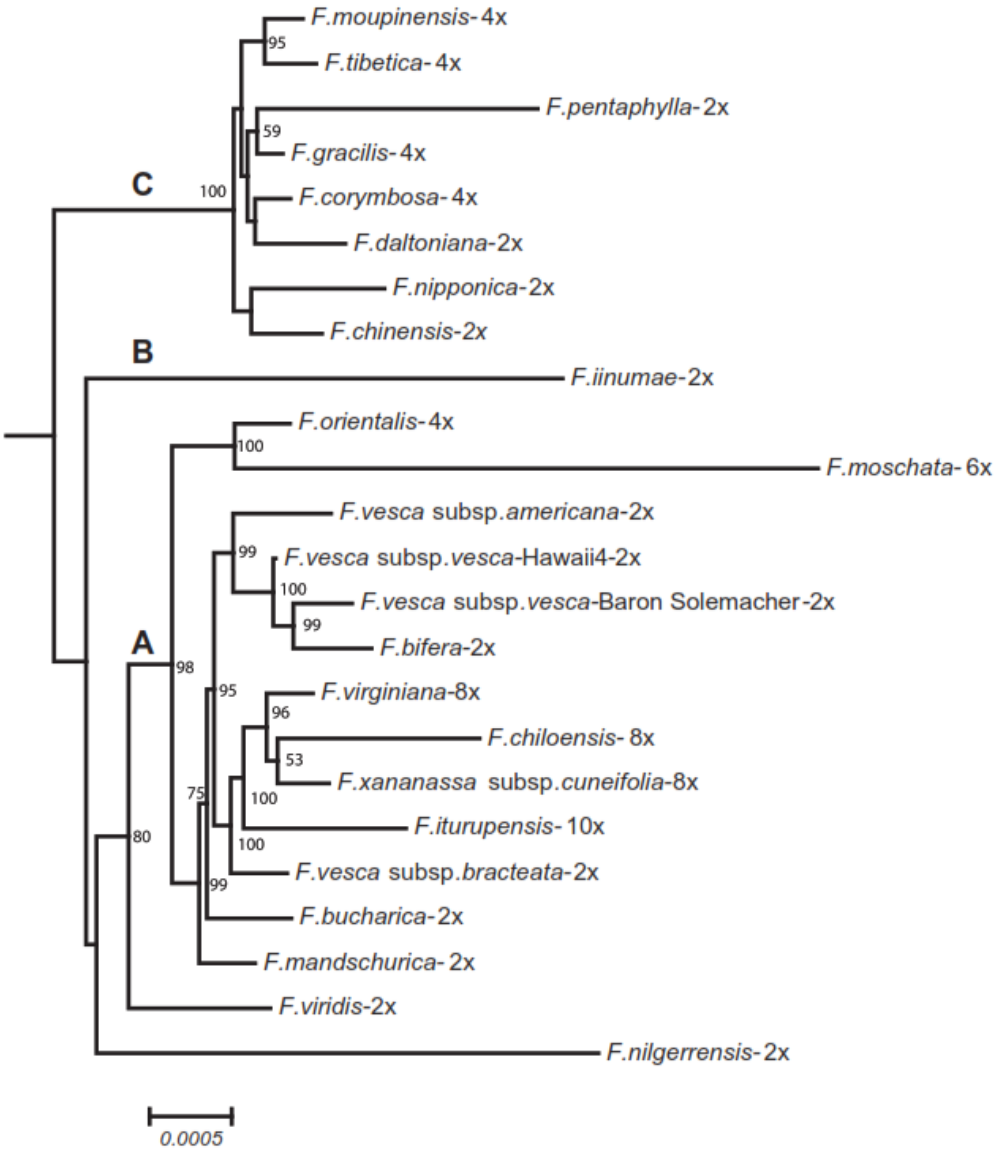


Figura 4. Relaciones filogenéticas dentro del género *Fragaria,* basada en el análisis de sus plastomas por el método de Maxima probabilidad (o ML, *Maximum Likelihood* en inglés). A, B y C son clados definidos por estudios filogenéticos previos. Mientras que los clados A (equivalente al clado "vesca") y el C (equivalente al clado "asiático") están bien definidos y reciben 98 y 100% de soporte mediante "bootstrap", el clado B (formado exclusivamente por *F. iinumae*) está separado de A y C, aunque con un soporte menor del 50%. Por otra parte, *F. nilgerrensis* y *F. viridis* no están incluidos en ninguno de los otros clados ni forman uno propio, si bien *F. viridis* parece estar relacionado con el clado A con un 80% de "boostrap". La cuarta especie que mantiene relaciones filogenéticas dudosas con el resto, *F. hayatai*, no está incluida en el análisis. Figura tomada de *Njuguna et al., 2013*.

Tabla 1. Especies y subespecies aceptadas del género Fragaria agrupadas filogenéticamente, mostrando su nivel de ploidía y distribución geográfica aproximada (Liston et al., 2014;Sobczyk, 2018;Folta and Barbey, 2019).

División	Especies	Ploidía	Distribución
Clado "vesca"	*F. vesca* L. ssp. *vesca*	2x	Eurasia
	F. vesca L. ssp. *bracteata*	2x	Norteamérica
	F. vesca L. ssp. *californica*	2x	Norteamérica
	F. vesca L. ssp. *americana*	2x	Norteamérica
	F. mandshurica	2x	Asia
	F. bucharica	2x	Himalaya
	F. orientalis	4x	Asia
	F. moschata	6x	Eurasia
	F. iturupensis	8x,10x	Isla Iturup (Kuriles)
	F. chiloensis	8x	Norteamérica, Chile, Argentina
	F. virginiana	8x	Norteamérica
	F. cascadensis	10x	Norteamérica
	F. x ananassa ssp. *ananassa (F. virginiana x F. chiloensis)*	8x	Mundial, especie cultivada
	F. × ananassa ssp. *cuneifolia (F. virginiana x F. chiloensis)*	8x	Norteamérica
	F. x bifera (F. vesca x F. viridis)	2x,3x	Europa
	F. x bringhurstii (F. vesca x F. chiloensis)	5x,6x,9x	California
Clado "asiático"	*F. daltoniana*	2x	Himalaya
	F. nipponica	2x	Japón
	F. nubicola	2x	Himalaya
	F. pentaphylla	2x	China
	F. chinensis	2x	China
	F. corymbosa	4x	China, Rusia
	F. gracilis	4x	China
	F. moupinensis	4x	China
	F. tibetica	4x	Tíbet
No resueltas	*F. hayatai*	2x	Taiwan
	F. iinumae	2x	Japón, Rusia
	F. nilgerrensis	2x	Asia
	F. viridis	2x	Eurasia

1.2. La fresa cultivada. Origen e importancia

La fresa cultivada, denominada comúnmente fresa, fresón o frutilla, es la fruta blanda de mayor importancia económica, con una producción mundial de más de 8,33 Mt, en la que España ocupa el sexto puesto en importancia, y ocupando una superficie de cultivo de unas 372.361 ha (FAOSTAT, 2018, Figura 5).

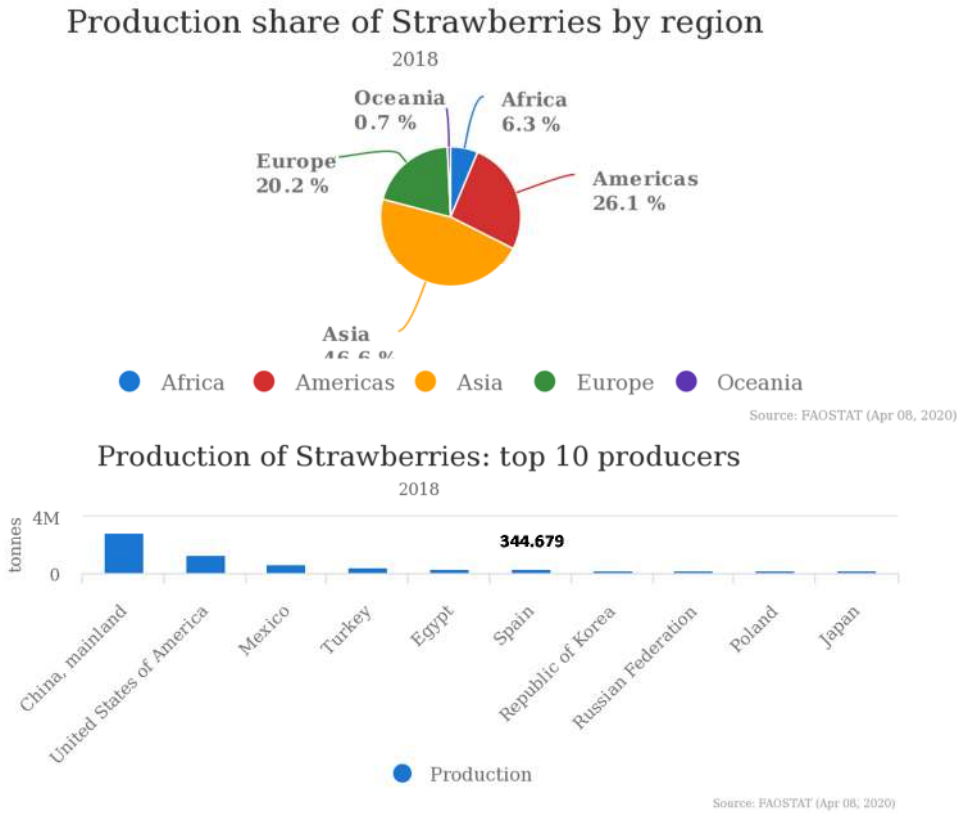


Figura 5. Datos sobre la producción mundial de fresa cultivada (FAOSTAT, 2018).

Históricamente, existen evidencias del consumo de fresas silvestres desde la edad de piedra en Europa, así como de la domesticación de *F. chiloensis* por indígenas sudamericanos hace más de 1000 años (Vergauwen and De Smet, 2019). En Europa, *F. vesca* ya fue cultivada en jardines romanos desde la antigüedad, mientras que otras especies, como *F. viridis* y *F. moschata* no se popularizaron hasta siglos más tarde, principalmente como especies ornamentales (Hummer and Hancock, 2009). Sin embargo, la fresa cultivada o fresón (*Fragaria × ananassa* ssp. *ananassa*) es el resultado de la hibridación fortuita entre dos especies americanas nativas: *F. chiloensis*, introducida en Europa en el siglo XVIII, y *F. virginiana*, que ya había sido importada del nuevo mundo en el siglo XVI. Debido a su gran tamaño, aroma y sabor, su cultivo se popularizo y extendió rápidamente por Europa y el resto del mundo. Pronto, los criadores europeos experimentaron con nuevos cruzamientos e híbridos, labor que ha continuado hasta nuestros días, dando lugar a los modernos cultivares existentes (Hummer and Hancock, 2009;Liston et al., 2014;Vergauwen and De Smet, 2019).

La fresa es muy apreciada por sus cualidades nutricionales y organolépticas. El fruto maduro se compone aproximadamente en un 90% de agua y en un 10% de sólidos solubles que incluyen numerosos compuestos de gran valor nutricional y propiedades saludables, principalmente antioxidantes y anti-inflamatorias (Hummer and Hancock, 2009;Fierascu et al., 2020). Entre ellos destacan los azúcares (principalmente glucosa, fructosa y sacarosa), los ácidos cítrico y ascórbico (vitamina C) y gran variedad de compuestos fenólicos, como las antocianinas. Como resultado de numerosos programas de mejora genética realizados por todo el mundo (aunque particularmente en Estados Unidos, Europa, China y Japón), el número de cultivares comerciales de fresa ha experimentado un continuo aumento. Sin embargo, las nuevas variedades tienden a sustituir a las generadas previamente, debido a la introducción de nuevos caracteres que proporcionan mejoras en: la adaptabilidad de los cultivos a determinadas regiones, producción y mejores cualidades nutricionales y organolépticas, entre otras (Hummer and Hancock, 2009;Mezzetti et al., 2018).

1.3. Genómica de la fresa diploide y octoploide

Durante las décadas pasadas, la mejora asistida por marcadores de la fresa comercial había estado experimentando un cierto retraso con respecto a otras especies cultivadas (Verma et al., 2018). Ello se ha debido, en buena medida, a la ausencia de genomas de referencia y a la complejidad genética de las especies poliploides, que dificultaba enormemente el análisis filogenético y los estudios comparativos acerca de sus orígenes y evolución. Esto no ha podido resolverse hasta muy recientemente, con la publicación de los genomas de referencia de *Fragaria vesca* y *Fragaria × ananassa* y herramientas moleculares para asistir a los programas de mejora genética y selección de la fresa cultivada (Edger et al., 2019;Hardigan et al., 2019;Whitaker et al., 2020).

1.3.1. Genómica de *Fragaria vesca*

No fue hasta 2011 que se dispuso del primer genoma de referencia, completamente ensamblado, de la especie diploide *F. vesca* ssp *vesca* “Hawaii 4” (Shulaev et al., 2011). Mediante la combinación de las plataformas Roche/454, Illumina/Solexa y Life Technologies/SOLID se obtuvo una primera versión del genoma (Fragaria vesca Whole Genome v1.0, o FvH4) formado por una secuencia de 198.1 Mb, alineada mediante un mapa de ligamiento de referencia y distribuida en 7 pseudocromosomas, conteniendo casi 35.000 *ab initio* genes, posteriormente depurados en la versión denominada “v1.0-hybrid”. Estos datos fueron puestos a disposición de la comunidad científica, principalmente en repositorios especializados como la Genome Database for Rosaceae (GDR), que es probablemente el principal recurso para la investigación en fresa y otras rosáceas (Jung et al., 2019). La disponibilidad del genoma de *Fragaria vesca* supone, desde ese momento, una alternativa ventajosa a los estudios genómicos y funcionales basados en la planta modelo *Arabidopsis thaliana*. En primer lugar, *Fragaria vesca*

ofrece ventajas similares, como son la de poseer un genoma relativamente pequeño (unas 240 Mb), rápido crecimiento (a pesar de ser una planta perenne), reproducción sexual, así como capacidad de multiplicación y propagación asexual y la existencia de sistemas de transformación genética eficientes. Por otra parte, se trata de una planta perteneciente a la familia Rosaceae, proporcionando la oportunidad de realizar complejos estudios evolutivos y funcionales dentro de este importante grupo de plantas eudicotiledóneas, muchos de cuyos miembros poseen alto valor económico.

Poco después de su publicación, el genoma fue actualizado introduciendo algunas correcciones (versión 1.1), aunque las anotaciones iniciales se mantuvieron. No es hasta 2015 que esta anotación es actualizada (Darwish et al., 2015), usando RNA-seq de hasta 25 tejidos diversos de *F. vesca* cv. YW5AF7, para obtener un nuevo transcriptoma de fresa (versión 1.1.a2).

Con posterioridad, el genoma es revisado comprobando la orientación y disposición de varios segmentos, usando para ello un mapa de ligamiento generado a partir de *F. vesca* ssp. *Bracteata (Tennessen et al., 2014)*. Esta nueva versión (Fvb v2.0.a1) supone una mejora sustancial en cuanto a la reorganización del ensamblado anterior, corrigiendo varias translocaciones, inversiones y localizaciones incorrectas de algunos segmentos en las versiones previas, además de añadir nuevas secuencias, previamente no localizadas, a los pseudocromosomas (208.9 Mb totales, de los cuales 207.0 Mb forman parte de los 7 pseudocromosomas). El nuevo ensamblado posee, además, un mayor grado de sintenia con otras especies relacionadas, como *Prunus*.

Aunque, en un primer momento, la versión 2 hereda las anotaciones generadas previamente (v1.0-hybrid y v1.1.a2), los avances producidos en las técnicas de mapeo y secuenciación de ARN (RNA-seq), particularmente la nueva Secuenciación SMRT (conocida más popularmente como PacBio) y la generalización de su uso posibilitan la aparición de una nueva versión de la anotación del genoma, la Fvb v2.0.a2. Esta nueva anotación combina la información de 100 librerías de RNA-seq, procedentes de diversos tejidos y supone una mejora muy significativa sobre las anteriores, ya que actualiza la anotación funcional, el mapeo y los modelos génicos (*gene models*) para más de 13.168 loci, incluye variantes alternativas de *splicing* para 7.370 genes y añade los extremos *5'*- y *3' –UTR* para 18.641 genes. Además, proporciona la anotación de más de 53.000 ARNs no codificantes, entre *miRNAs*, *lncRNAs* y *small RNAs*.

El último y, hasta ahora, más importante avance en la obtención de un genoma de referencia de alta calidad se ha producido recientemente, usando la secuenciación SRMT para producir un nuevo genoma de referencia denominado *F. vesca* v4 (Edger et al., 2018) y anotación v4.0.a1. El nuevo genoma de referencia añade cerca de 25 Mb de nuevas secuencias, incluyendo extremos teloméricos para los 7 pseudocromosomas y mejorando la contigüidad y orientación de los fragmentos secuenciados gracias a su mayor longitud media (N50), de hasta 300 veces superior a la secuenciación original. Una comparación entre las versiones v2.0.a1 y v4.0.a1 pone de manifiesto la existencia de errores de ensamblado y regiones no mapeadas en la primera (Figura 6).

Poco después, es publicada una nueva anotación denominada v4.0.a2 (Li et al., 2019), basada en el uso de los datos transcriptómicos previos generados mediante RNA-seq, del mismo modo que ocurriera con la anterior versión v2.0.a2. De nuevo, en esta última versión se describen los modelos génicos completos para más de 34.000 genes, variantes de *splicing* alternativo, incluyendo los extremos *UTR* y ARNs no codificantes. La importancia de estos avances no sólo radica en el hecho de proporcionar una secuencia más completa y anotaciones génicas más exactas, mejorando así, por ejemplo, los estudios de expresión génica. También posibilita la realización de estudios filogenéticos y evolutivos más detallados, tanto dentro del género *Fragaria* como en sus relaciones con otras especies.

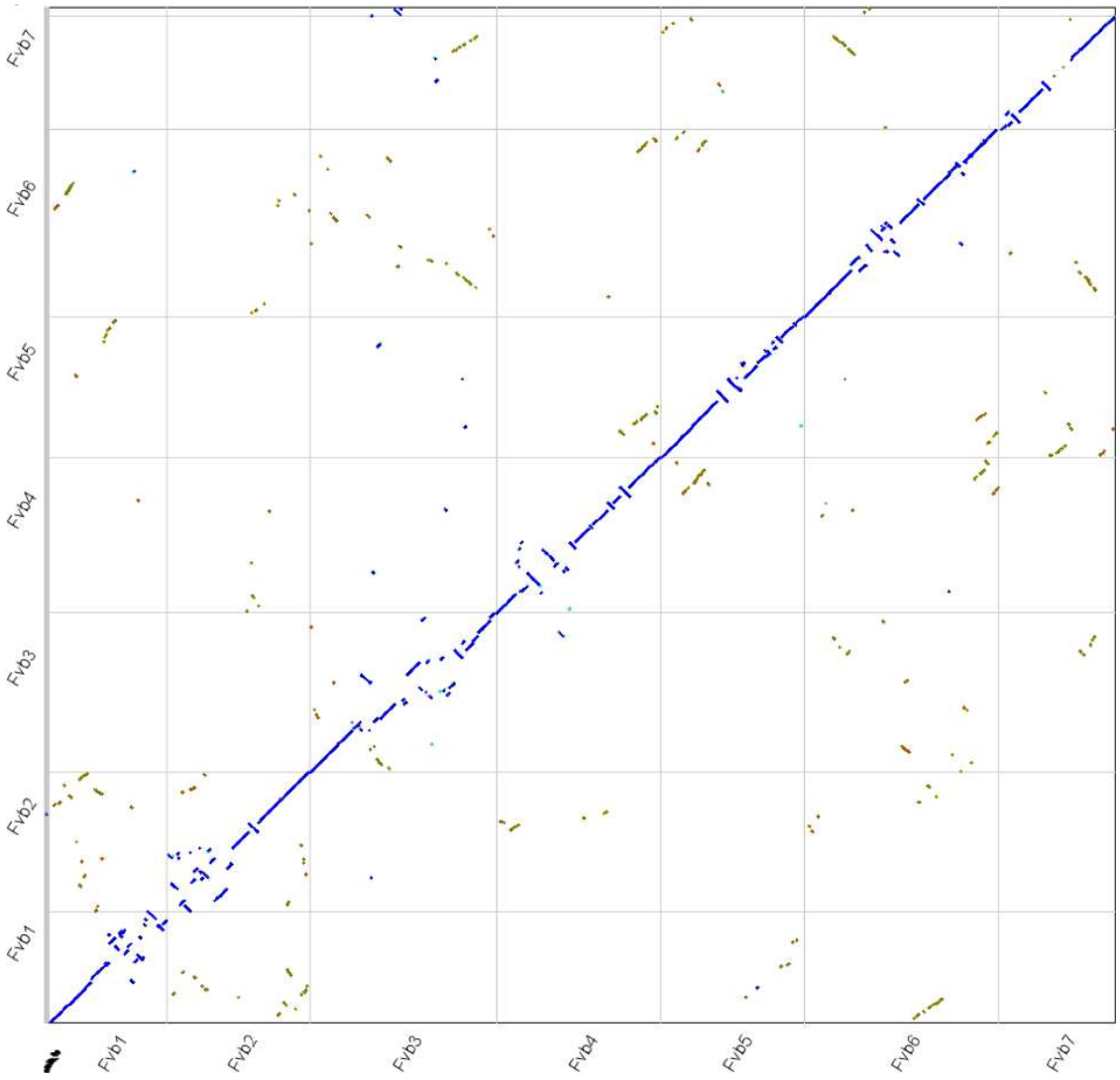


Figura 6. Macrosintenia entre los genomas v2.0.a1 (eje y) y v4.0.a1 (eje x) de *Fragaria vesca.* Las regiones equivalentes del genoma están coloreadas en azul. Algunas de ellas muestran una colinearidad invertida, mientras que otras están colocadas incorrectamente o ausentes (puntos o líneas emplazadas fuera de la diagonal y huecos en la misma, respectivamente). Figura tomada de Edger et al. (2018).

1.3.2. Genómica de *Fragaria x ananassa*

La naturaleza alógama y alopoliploide del fresón ha dificultado considerablemente el estudio del genoma y sus orígenes (Hirakawa et al., 2014). Los primeros estudios, limitados a análisis citológicos clásicos de configuraciones durante la meiosis proponen las siguientes fórmulas para la composición cromosómica del genoma de *Fragaria x* ananassa: AABBBBCC (tres subgenomas diferentes) y, posteriormente, AAA'A'BBBB o AAA'A'BBB'B' (Folta and Davis, 2006). Estos modelos implican la existencia de tres o de cuatro orígenes diploides diferentes contribuyendo al genoma, por tanto, alopoliploide, de la fresa cultivada, siendo uno de ellos identificado como perteneciente a *F. vesca.* Estudios filogenéticos posteriores (Tennessen et al., 2014) proponen la participación de *F. vesca* (Av), *F. iinumae* (Bi) y un autotetraploide relacionado con *F. iinumae* (B1, B2) para dar lugar al alo-octoploide, que tendría la fórmula cromosómica 2Av, 2Bi, 2B1, 2B2. Sin embargo, aunque la participación de *F. vesca* y *F. iinumae* es confirmada tanto por el análisis genético de los genes nucleares como de los plastomas, otros trabajos no apoyan el origen de los subgenomas B1 y B2 como procedentes de un autotetraploide relacionado con *F. iinumae* y apuntan a la participación de uno o dos diploides desconocidos, pero relacionados con *F. viridis*, *F. bucharica* o *F. mandshurica* (Sobczyk, 2018). Esta controversia no pudo ser completamente resuelta en un primer proyecto de secuenciación del genoma de *Fragaria x ananassa* y cuatro especies diploides: *F. iinumae*, *F. nipponica*, *F. nubicola* y *F. orientalis*, además de *F. vesca*, previamente secuenciada (Hirakawa et al., 2014). Se obtuvo un ensamblado parcial del genoma octoploide que sólo permitió confirmar una participación mayoritaria de secuencias pertenecientes a *F. vesca* y *F. iinumae*, mientras que no pudo determinarse el origen de alrededor del 25% restante.

Con la reciente publicación del genoma completo y anotación de *Fragaria × ananassa* cv Camarosa (Edger et al., 2019), se identifica a *F. vesca*, *F. iinumae*, *F. viridis* y *F. nipponica* como los donantes diploides de los diferentes subgenomas encontrados en la fresa cultivada. Según la hipótesis de Edger y colaboradores, las especies asiáticas habrían hibridado entre sí para formar primero un intermediario tetraploide (*F. nipponica* x *F. iinumae*), que a su vez habría formado un híbrido hexaploide, similar a *F. moschata*, incorporando el genoma de *F. viridis*. Finalmente, tras cruzar posiblemente el estrecho de Bering, el hexaploide se habría introducido en el continente americano e hibridado con *F. vesca* ssp. *bracheata*, para dar lugar, con el tiempo, a los progenitores octoploides de la fresa cultivada, *F. chiloensis* y *F. virginiana* (Figura 7). Este trabajo revela, además, que el subgenoma derivado de *F. vesca* es expresado de forma dominante y contribuye mayoritariamente en aspectos clave de la fisiología de la planta, como los mecanismos de resistencia frente a patógenos y las rutas metabólicas responsables de las características organolépticas del fruto y que le dan su aroma, color y sabor. Además, mientras que los genes procedentes de los otros progenitores han sufrido un mayor grado de fraccionamiento génico (esto es, pérdida de genes homeólogos tras un evento de poliploidización), los pertenecientes al subgenoma de *F. vesca* habrían sido favorecidos por la presión de selección evolutiva, siendo retenidos preferencialmente e, incluso, sustituyendo a genes de otros

subgenomas mediante recombinación homeóloga. El fraccionamiento sesgado habría producido que los subgenomas procedentes de *F. viridis* y *F. nipponica* se hayan fraccionado en mayor medida tras la hibridación y constituyan, en la actualidad, meros fragmentos de los cromosomas originales. En cambio, aquellos procedentes de *F. vesca* y *F. iinumae* habrían experimentado menores alteraciones y han sido retenidos en los híbridos en mayor medida, particularmente el primero. Este hecho explicaría los resultados que condujeron a las anteriores hipótesis sobre los orígenes de los genomas octoploides.

Sin embargo, se han presentado algunas objeciones a la metodología usada en este trabajo que ponen en duda la composición subgenómica de la fresa octoploide, resucitando la hipótesis de la participación de *F. iinumae* en tres de los subgenomas octoploides a través de un híbrido autotetraploide (Liston et al., 2020). Por otra parte, se ha propuesto que el género *Fragaria* es especialmente propenso a las hibridaciones interespecíficas, por lo que la presencia de secuencias de *F. viridis* en el octoploide podría ser resultado del flujo genético con otras especies diploides, como *F. iinumae* (Feng et al., 2020). Por tanto, aunque se presentaron evidencias adicionales para apoyar la hipótesis de los cuatro progenitores diploides diferentes para *Fragaria × ananassa* (Edger et al., 2020), el debate continua abierto y, posiblemente, serán necesarios datos adicionales y análisis más detallados para confirmar una u otra hipótesis.

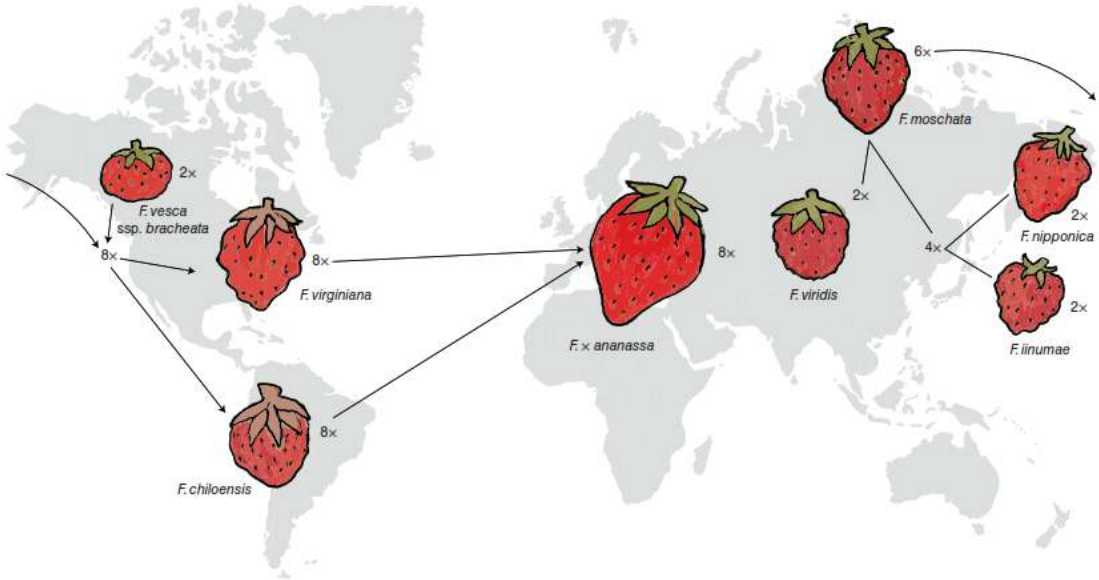


Figura 7. Reconstrucción de los eventos de hibridación entre distintas especies de *Fragaria,* distribuidas entre los continentes euroasiático y americano, que dieron lugar a las especies octoploides progenitoras de *Fragaria x ananassa*. Figura tomada de Bertioli (2019) (https://doi.org/10.1038/s41588-019-0365-3)

1.4. Enfermedades de la fresa y mecanismos de defensa

1.4.1. Principales enfermedades de la fresa cultivada

La planta de fresa, así como sus frutos, se ven expuestos a un gran número de plagas y enfermedades, tanto durante las etapas de crecimiento y desarrollo vegetativo que culminan en la formación y maduración de los frutos, como en la postcosecha, una vez recogido el fruto y siendo preparado para su entrada en los canales comerciales. Entre sus principales plagas se encuentran artrópodos y otros invertebrados, siendo los más comunes distintas especies de ácaros, pulgones, trips, nematodos, caracoles y babosas. Estos pequeños animales pueden, a su vez, actuar como vectores para virus, fitoplasmas y otros microorganismos dañinos o facilitar la colonización de los tejidos afectados por, particularmente, bacterias y hongos causantes de distintos tipos de enfermedades de la planta y del fruto (Maas, 2004).

Se conoce un gran número de virus que usan como huésped a diferentes especies de la familia Rosaceae (http://bio-mirror.im.ac.cn/mirrors/pvo/vide/famly114.htm), incluyendo más de 20 especies en el género *Fragaria* (Martin and Tzanetakis, 2013;Nellist, 2018) y que son transmitidos, frecuentemente, por nematodos e insectos. El principal efecto de las enfermedades víricas en el cultivo de la fresa es una importante reducción de la producción en las plantas afectadas, causada por un declive en la vitalidad de las mismas que puede producir, en casos extremos, la pérdida total de la planta. El control de las virosis en la fresa, como en otras plantas de interés agronómico, se centra en la siembra de plantas procedentes de vivero certificadas libres de virus, en mantener unas buenas prácticas culturales y en el control de los vectores de transmisión.

Las enfermedades causadas por bacterias y hongos son, sin embargo, las de mayor incidencia e importancia económica en el cultivo de la fresa, afectando tanto a las partes vegetativas como a los frutos (Maas, 2004;Nellist, 2018). Existen pocos patógenos bacterianos descritos como agentes causales de enfermedades en la fresa. De ellos, el de mayor impacto económico es *Xanthomonas fragariae*, un patógeno distribuido globalmente, muy resistente a condiciones ambientales y a productos fitosanitarios. Esta bacteria es un patógeno específico del género *Fragaria* y causa la enfermedad conocida como “mancha aceitosa” en las hojas (ALS, *angular leaf spot*).

Por el contrario, numerosas especies de hongos patógenos pueden producir enfermedades tanto de la planta como del fruto de fresa, siendo estas las de mayor importancia, tanto por su variedad como por su incidencia en distintas regiones productoras, siendo responsables de cuantiosas pérdidas económicas (Maas, 2004;Nellist, 2018). En la Tabla 2 se recogen las principales enfermedades fúngicas que afectan al cultivo de la fresa, así como las especies de hongos causantes y las regiones geográficas más afectadas.

Tabla 1. Principales enfermedades fúngicas de la fresa (planta y fruto). Se indican las regiones geográficas con mayor incidencia de las mismas. Adaptado de Nellist (2018).

Enfermedad	Hongo causante	Región geográfica afectada			
		Europa	USA	Asia	Australia
Moho gris	*Botrytis cinerea*	X	X	X	X
Antracnosis	*Colletotrichum acutatum*	X	X		
	Colletotrichum gloeosporioides		X	X	
Marchitez vascular	*Fusarium oxysporum* spp. *fragariae*		X	X	X
Mancha de las hojas	*Gnomonia comari*	X			X
Mancha negra de las hojas	*Alternaria alternata*	X		X	
Viruela	*Mycosphaerella fragariae*		X		
Pudrición negra de la raíz	*Pythium* spp., *Fusarium* spp., *Rhizoctonia* spp.		X		X
Pudrición de la corona	*Phytophthora cactorum*	X	X		
Pudrición roja de la raíz	*Phytophthora fragariae*	X	X		
Moho polvoso	*Podosphaera aphanis*	X		X	
Pudrición blanda del fruto	*Rhizopus* spp.	X	X		
	Mucor spp.	X			
Verticilosis	*Verticillium dahliae*	X	X		

1.4.2. Antracnosis en fresa

La antracnosis es una de las principales enfermedades de la fresa cultivada, extendida globalmente y que afecta a todos los principales países productores. Está causada por hongos ascomicetos del género *Colletotrichum*, considerado uno de los principales hongos fitopatógenos (Dean et al., 2012). El género cuenta con más de 190 especies, causantes de la enfermedad en multitud de especies monocotiledóneas y dicotiledóneas de gran importancia económica (por ejemplo, maíz, trigo, manzana, aguacate, naranja, almendra, etc.) (Crouch et al., 2014;Jayawardena, 2016).

La planta de fresa se ve afectada, principalmente, por tres especies: *C. acutatum, C. gloeosporioides* y *C. fragariae*. Aunque las tres especies pueden causar lesiones en cualquier parte de la planta, *C. acutatum* es el principal causante de las lesiones en los frutos, mientras que *C. gloeosporioides* afecta, principalmente, a la corona, peciolos y

estolones (Maas, 2004;Peres et al., 2005;Nellist, 2018). El principal síntoma de la enfermedad es la aparición de lesiones necróticas exteriores de apariencia reseca, oscuras y algo alargadas, de color grisáceo a negro y ligeramente hundidas en frutos, hojas, estolones y peciolos, mientras que en las coronas aparecen zonas de color rojizo en los tejidos internos (Maas, 2004) (Figura 8).

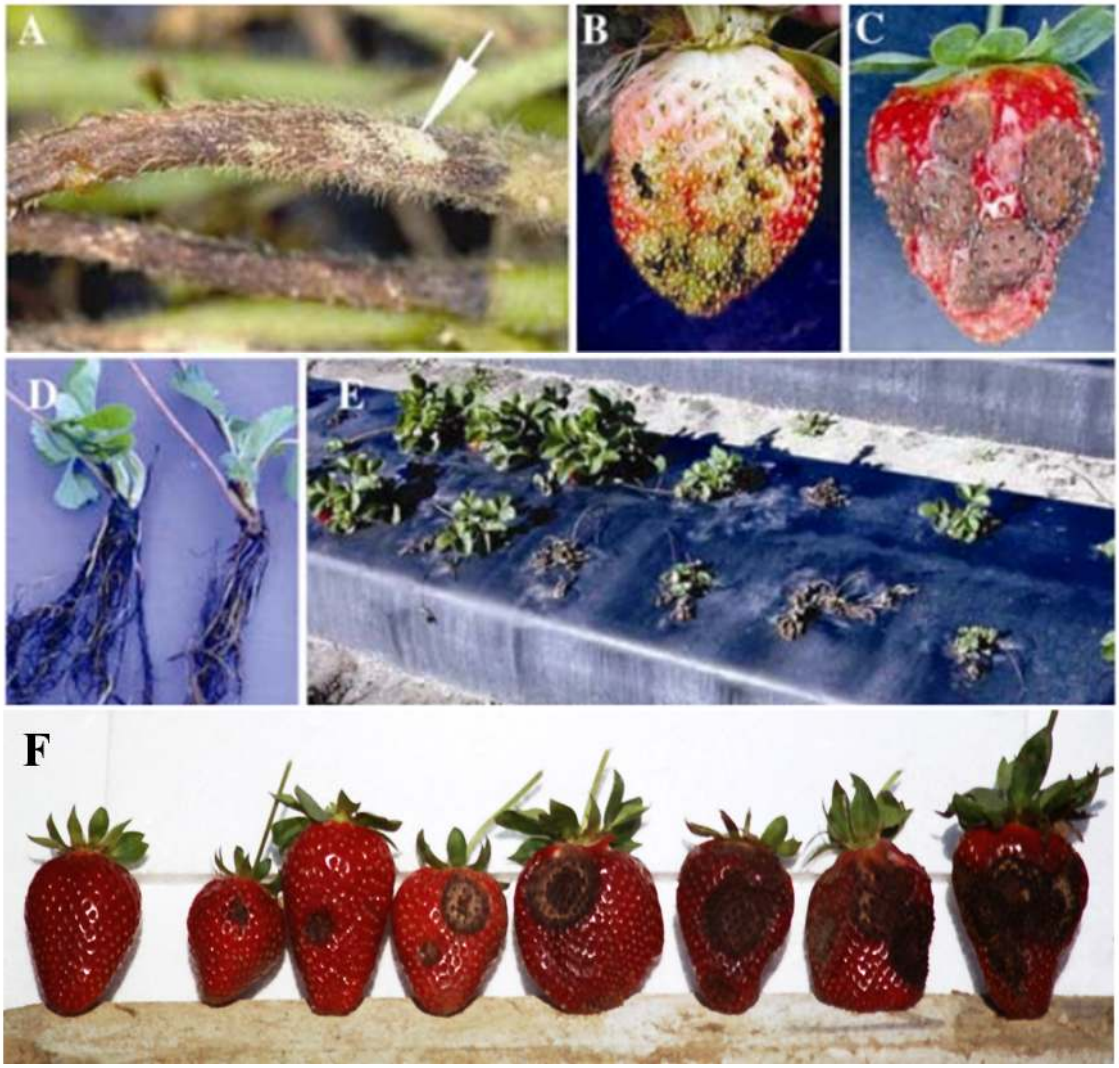


Figura 8. Sintomatología de la antracnosis en fresa, causada por *Colletotrichum acutatum*, infectando peciolos (A, con detalle de acérvulo - flecha - en la superficie de la lesión); fruto inmaduro (B) y maduro (C); y raíces (D), que adquieren un color exterior negruzco causado por la necrosis interna. Las plantas pueden, finalmente, marchitarse y morir cuando corona y raíces son afectadas severamente por la enfermedad (E). Los frutos pueden presentar distintos grados de lesiones, que evolucionan durante varios días, y sobre las que se desarrollan acérvulos (F). Figura adaptada de Maas (2004) (paneles A-E) y fotografía (F) de una infección natural de *C. acutatum* en frutos de la variedad Camarosa (José L. Caballero, 2008).

1.4.3. Desarrollo de la enfermedad y respuesta de defensa

El ciclo de vida de *Colletotrichum* spp. comienza con la germinación de las conidias sobre la planta huésped, formándose rápidamente una estructura de fijación denominada apresorio, desde la cual se emiten hifas que penetran la cutícula de las células epidérmicas del huésped y colonizan los tejidos internos, alimentándose de las células vivas mediante haustorios (Curry et al., 2002;Horowitz et al., 2002;Peres et al., 2005). Esta fase (Figura 9), en la que *Colletotrichum* exhibe un comportamiento biotrofo, puede tener una duración variable, en torno a 24 horas en el caso de la interacción con la fresa, aunque según las condiciones, especies o cultivares resistentes, pueden desarrollarse infecciones quiescentes relativamente prolongadas, con crecimiento micelial sobre la superficie y formación de conidias secundarias (Curry et al., 2002). El hongo sortea la respuesta de defensa de la planta (Jones and Dangl, 2006), burlando la denominada defensa basal o PAMP-Triggered Immunity (PTI, inmunidad inducida por PAMP) mediante la acción de efectores que producen la susceptibilidad del huésped (ETS, *Effector-Triggered Susceptibility*). Se conocen algunos mecanismos que el hongo utiliza evitar su detección por parte de los receptores (PRRs, *Pattern Recognition Receptors*) de la planta, como la deacetilación de la quitina de las hifas para formar quitosano (Munch et al., 2008) o la secreción de proteínas de unión a quitina (Takahara et al., 2016), lo que evita la degradación de la quitina por parte de las quitinasas de planta y la consiguiente generación de fragmentos de pared celular del hongo, que actuarían como elicitores de las defensas de la planta. Por otra parte, el hongo también produce efectores que suprimen los mecanismos de la defensa basal, particularmente los relacionados con la respuesta hipersensible (HR) que producen muerte celular y tratan de formar una barrera a la penetración del patógeno (O'Connell et al., 2012;Irieda et al., 2014;de Queiroz et al., 2019;Tomas-Grau et al., 2019).

A la fase biotrofa le sigue una fase de crecimiento necrotrofo, en la que el hongo se extiende y coloniza masivamente las células y tejidos del huésped, provocando la necrosis que se traduce en los síntomas visibles antes mencionados (Figura 8). En esta fase, la planta desarrolla una nueva fase de respuesta de defensa denominada ETI (*Effector-Triggered Immunity*), en la que proteínas específicas del patógeno son reconocidas por receptores de la planta (Jones and Dangl, 2006), particularmente por proteínas pertenecientes a la superfamilia NB-LRR (Lee and Yeom, 2015;van Eck and Bradeen, 2018), codificados por los llamados genes-R (R por resistencia). La consiguiente transducción de la señal desencadena cambios transcripcionales en las células que activan las rutas de defensa y que desemboca, normalmente, en la activación de la respuesta HR. Al contrario que en la fase biotrofa, el hongo secreta factores de virulencia que estimulan, e incluso favorecen, la respuesta HR en los tejidos infectados (Mengiste, 2012).

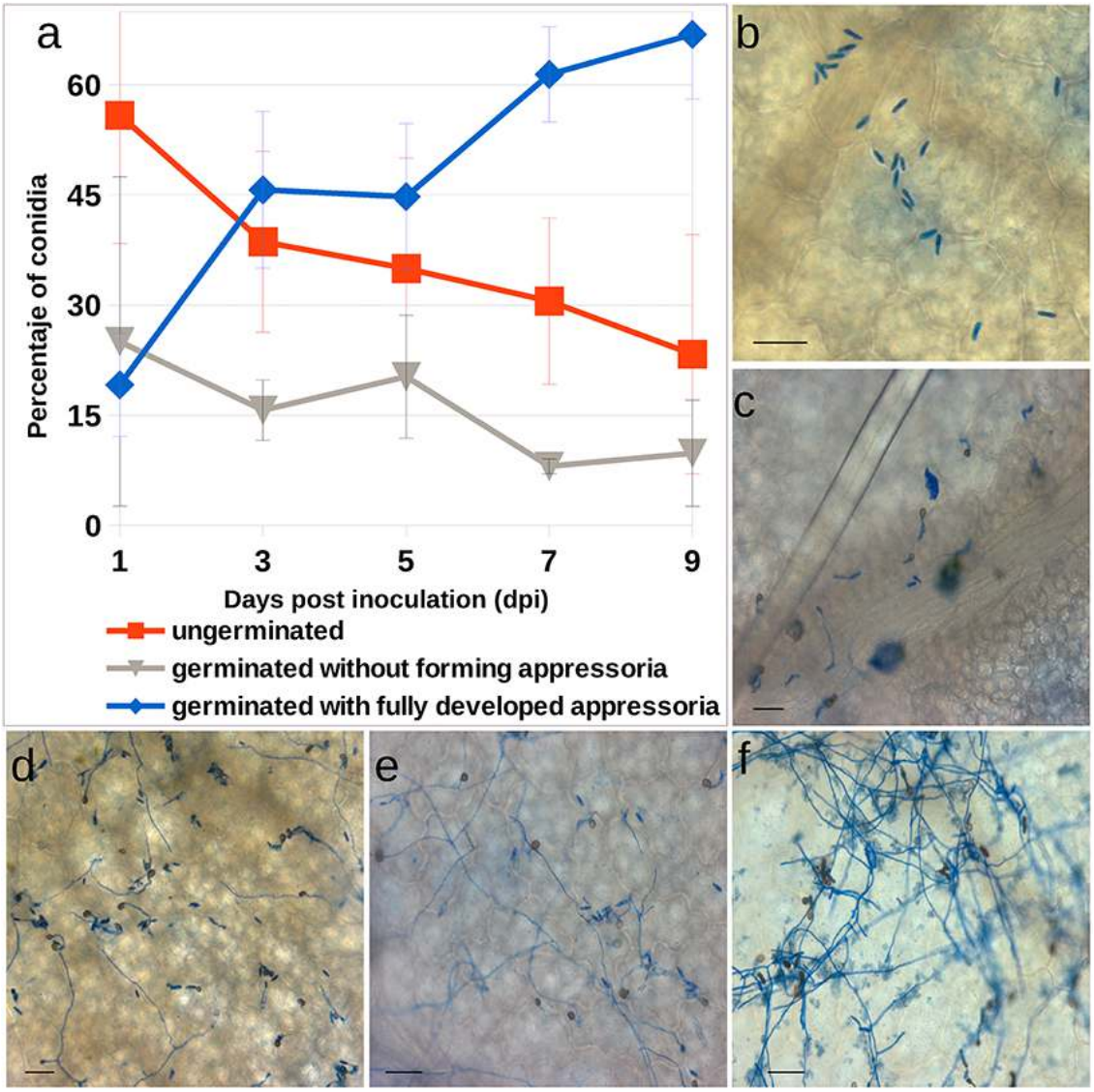


Figura 9. Desarrollo de la fase biotrofa de *Colletotrichum acutatum* sobre discos de hoja de fresa (Camarosa), mostrando un incremento de la germinación de las conidias durante varios días tras la inoculación **(a-f)**. Las conidias germinadas forman apresorios que se melanizan rápidamente **(c)**, formándose un micelio superficial que incrementa en extensión tras varios días **(d,e,f)** y puede formar conidias secundarias **(f)**. Tomado de Amil Ruiz et al. (2016).

Así, el hongo secreta amonio, que produce una alcalinización de los tejidos invadidos y aumenta la virulencia de la infección (Prusky et al., 2001;Peres et al., 2005;Alkan et al., 2015), así como enzimas líticas que degradan los tejidos afectados y efectores que modifican o anulan las defensas efectivas del huésped y detoxifican las sustancias tóxicas para el hongo, particularmente las especies reactivas de oxígeno (ROS) (Amil-Ruiz et al., 2016;Zhang et al., 2016;Liang et al., 2018;Zhang et al., 2018;He et al., 2019;Tomas-Grau et al., 2019). Finalmente, el hongo forma acérvulos, estructuras productoras de conidias, en las superficies infectadas, dispersándose mediante el viento o gotas de humedad producidas por el riego o la climatología.

Al alternar fases biotrófica y necrotrófica en su ciclo de vida, *Colletotrichum* spp. es considerado un patógeno hemibiotrofo. Sin embargo, el ciclo de la enfermedad puede presentar características variables según el huésped y el tejido infectado (Peres et al., 2005). Por ejemplo, *C. acutatum* se comporta esencialmente como biotrofo o necrotrofo en especies de *Citrus* según se desarrolle en hojas o en pétalos, respectivamente. En el caso de la interacción de *C. acutatum* con la planta de fresa, particularmente en frutos maduros, la fase biotrofa es tan breve que el patógeno tiene un desarrollo fundamentalmente necrotrofo, por lo que se ha especulado que este comportamiento podría ser más bien una modificación de la necrotrofía que una auténtica hemibiotrofía (Curry et al., 2002). La infección permanece quiescente en frutos no maduros, observándose, sin embargo, la germinación de las esporas con formación de apresorios melanizados. No es hasta el estadio de fruto rojo (maduro) que el hongo inicia la fase de crecimiento necrotrófica (Guidarelli et al., 2011). Este comportamiento se debe probablemente a la prevalencia de barreras mecánicas (cutícula y paredes celulares) y químicas (principalmente compuestos fenólicos) en los frutos inmaduros frente a los maduros (Amil-Ruiz et al., 2011). Durante la maduración, el fruto experimenta una pérdida de firmeza debido a la despolimerización de componentes de la pared celular y solubilización de pectinas (Molina-Hidalgo et al., 2013;Paniagua et al., 2017), así como profundos cambios metabólicos que, entre otros, producen un incremento de azúcares y modificaciones en las rutas de biosíntesis de flavonoides y antocianinas para producir compuestos relacionados con el color y el aroma del fruto (Amil-Ruiz et al., 2011;Medina-Puche et al., 2016;Wang et al., 2020). Estos cambios, en conjunto, contribuyen a un aumento de la susceptibilidad del fruto maduro a *C. acutatum* (Guidarelli et al., 2011). Además, recientemente se ha comprobado que la expresión global de genes R decae durante el proceso de maduración, reduciéndose aproximadamente a la mitad en frutos maduros (Barbey et al., 2019). Los mecanismos de defensa desplegados por las células de fresa conllevan cambios a nivel transcripcional que remodelan la fisiología de los tejidos infectados de la planta y la preparan para repeler el ataque del patógeno. Varios trabajos han caracterizado, a nivel molecular, los cambios transcriptómicos producidos en diversos tejidos de fresa a consecuencia de la infección por *Colletotrichum* spp. (Casado-Díaz et al., 2006;Guidarelli et al., 2011;Guidarelli et al., 2014;Amil-Ruiz et al., 2016;Zhang et al., 2018). Estos estudios han revelado cambios en varios procesos biológicos, alterados a consecuencia de la infección y que producen, principalmente:

1. Incremento de la síntesis y fortalecimiento de las paredes celulares.
2. Cambios en la expresión de PRRs y genes relacionados con la percepción del patógeno y la transducción de la señal (genes-R y quinasas, principalmente). Notablemente, la expresión de algunos genes-R es reprimida por la infección, mientras que otros son inducidos por la misma.
3. Expresión de genes implicados en la transcripción, síntesis y secreción de nuevas proteínas: genes de ARN ribosómico, helicasas y factores de transcripción (WRKY, NAC, ERF entre otros)

4. Activación de la respuesta de defensa sistémica (SAR, Systemic Acquired Resistance), con expresión de genes típicos de defensa: proteínas PR, pertenecientes a varios grupos (quitinasas, glucanasas, etc), así como de genes relacionadas con la síntesis y señalización de las principales rutas reguladoras de la defensa en plantas, mediadas por las fitohormonas ácido salicílico (SA) y ácido jasmónico (JA). Particularmente, la primera.

En general, la ruta de defensa mediada por SA es efectiva típicamente (pero no exclusiva) en las respuestas de defensa frente a patógenos biotrofos y hemibiotrofos, mientras que la mediada por JA (también denominada ruta de JA/ET, ya que la fitohormona etileno –ET- participa sinérgicamente con el JA) es activada frente a necrotrofos (Robert-Seilaniantz et al., 2011;Pieterse et al., 2012). Ambas rutas interaccionan antagónicamente y son finamente reguladas, de forma que el tipo de respuesta desplegada por la planta frente a los distintos tipos de patógenos (biotrofos, necrotrofos o hemibiotrofos) depende de un delicado balance entre ellas (Mengiste, 2012;Pieterse et al., 2012;Yang et al., 2019).

Los estudios demuestran que la activación de los mecanismos de defensa es, sin embargo, incompleta (Figura 10) (Amil-Ruiz et al., 2016). A pesar de que los niveles de SA y JA medidos en tejidos infectados por *C. acutatum* se incrementaron según lo esperado, no se observaron cambios significativos en la expresión de ciertos genes, comúnmente usados como marcadores moleculares característicos de cada ruta. Así, la expresión de genes como *FaGST* y *FaPR1.1*, que experimentan un incremento característico en respuesta a la producción de SA por parte de las plantas; y *FaPDF1*, *FaLOX2-1* y *FaJAR1*, que hacen lo propio en el caso del JA; no experimentó cambio alguno durante la infección en corona a pesar de que otros muchos genes clave experimentaron un aumento de su expresión según lo esperado. En particular, los genes codificantes de factores de transcripción *FaWRKY70* y *FaWRKY33*, denominados así por ser homólogos a aquellos de *Arabidopsis thaliana* y que participan en la regulación cruzada de las rutas de SA y JA respectivamente, incrementaron su expresión de manera continua. Estos resultados fueron los primeros en sugerir que el control represivo de FaWRKY33 sobre la ruta de SA no funciona adecuadamente y es solo parcial en los tejidos de fresa infectados por *C. acutatum*, lo que puede conducir a un antagonismo con la ruta de JA, haciendo que FaWRKY70 bloquee la expresión de aquellos genes de defensa de respuesta a JA de mayor eficacia en la planta para evitar el desarrollo necrotrofo del hongo, de forma que esta ruta de JA tampoco transcurre de forma efectiva y el patógeno puede progresar en su desarrollo e infección (Figura 10). De hecho, la manipulación de las rutas de señalización mediadas por SA y JA es una estrategia empleada habitualmente por bacterias y hongos patógenos para burlar las defensas del huésped (Rahman et al., 2012;Zhang et al., 2017).

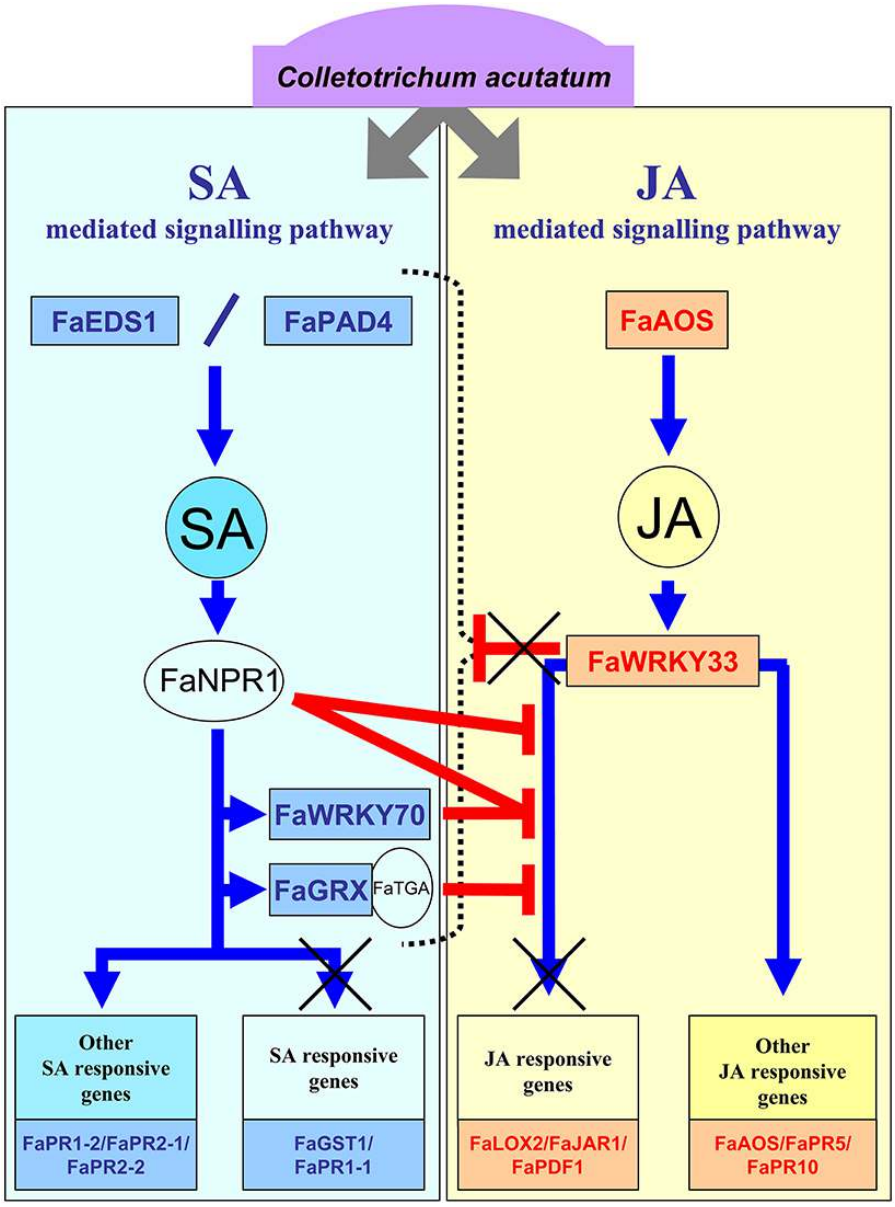


Figura 10. Modelo de activación incompleta de de las rutas de defensa mediadas por SA y JA en fresa tras la infección por *C. acutatum* (ver explicación en el texto). Figura tomada de Amil-Ruíz et al. (2016).

Estos resultados han sido corroborados y ampliados por trabajos posteriores sobre la interacción fresa-*C. fructicola* (Zhang et al., 2018;He et al., 2019). En ellos se muestra que la inducción de la respuesta de defensa de la fresa es insuficiente debido a alteraciones observadas en las rutas de SA y JA/ET. Además, la intervención de las rutas de señalización mediadas por las hormonas ácido abscísico (ABA) y auxinas podrían intervenir en el balance global de la defensa. (Zhang et al., 2018). Ciertamente, es sabido que estas fitohormonas pueden actuar de manera antagónica con las rutas de defensa mediadas por JA y SA, respectivamente (Robert-Seilaniantz et al., 2011), por lo que su activación durante la respuesta de defensa podría incrementar la susceptibilidad de la fresa frente al patógeno. Por tanto, la principal estrategia del

hongo sería la de manipular la homeostasis de SA en los tejidos infectados. En un primer momento, mediante la degradación de SA y la inhibición de su síntesis, de lo cual se encargarían un conjunto de efectores, formado por siete genes del hongo, que codifican tres tipos de efectores diferentes según su actividad: isocorismatasa y corismato mutasa, que retiran el isocorismato y corismato (precursores del SA) combirtiéndolo en otros metabolitos; y salicilato hidroxilasa, que degrada directamente el SA. Posteriormente, la infección podría modificar la respuesta de defensa mediada por SA mediante la activación de las rutas de señalización mediadas por JA y ABA (He et al., 2019).

Hoy día no existen cultivares de fresa verdaderamente resistentes a la antracnosis. En programas de mejora se han identificado algunos marcadores moleculares y QTLs (*Quantitative Trait Loci*) asociados a germoplasma de fresa menos susceptible a *C. acutatum* y *C. gloeosporioides* (Denoyes-Rothan et al., 2005;Lerceteau-Kohler et al., 2005;Anciro et al., 2018;Salinas et al., 2019) que corresponden, esencialmente, a regiones que codifican genes R (Barbey et al., 2019). Así, en estos cultivares, se ha observado una reducción en la germinación de las conidias y un retraso de varios días en la aparición de los síntomas. Asímismo, se han encontrado diferencias transcriptómicas cualitativas y cuantitativas evidentes entre los cultivares de fresa con distinto grado de susceptibilidad, tras ser infectados por *Colletotrichum* spp., lo que apoya que los mecanismos de resistencia son poligénicos (Casado-Díaz et al., 2006;Amil-Ruiz, 2013;Zhang et al., 2016;Wang et al., 2017). Por tanto, la mejora genética de la resistencia frente a *Colletotrichum* spp. en fresa se ve dificultada por tratarse de caracteres cuantitativos, que se expresan de forma diferente en distintos órganos de la planta y que pueden heredarse independientemente (Jacobs et al., 2019). No obstante, la resistencia adquirida mediante programas de mejora no suele ser duradera a largo plazo en la mayoría de las plantas, ya que los patógenos evolucionan para adaptarse y burlar las nuevas defensas, ya sean caracteres cuantitativos o monogénicos (Lo Iacono et al., 2013;Cowger and Brown, 2019).

Actualmente, los principales mecanismos de control de la enfermedad en fresa son la prevención de su aparición y diseminación, mediante el uso de buenas prácticas culturales, o el empleo de fungicidas para tratar el material infectado (Maas, 2004). Es, por tanto, necesario y de gran importancia conocer no solo los componentes moleculares que intervienen en la defensa de fresa a patógenos, sino comprender los complejos mecanismos de regulación de las diferentes rutas de defensa participantes en cada caso (patógeno). Ello permitirá la identificación de genes potencialmente útiles en futuros programas de mejora genética dirigidos a conseguir cultivares que exhiban una mayor y más duradera resistencia (Nellist, 2018;Cowger and Brown, 2019). Consecuentemente, esto favorecerá la reducción del uso de fungicidas permitiendo el desarrollo de variedades más respetuosas con el medio ambiente y con la salud pública.

1.4. Factores de transcripción WRKY y proteínas VQ

Los factores de transcripción (FT) son reguladores de la transcripción, que actúan en conjunto con otros componentes de la maquinaria transcripcional, modulando la expresión de genes diana en aquellas células y tejidos en las que se requiere una expresión génica determinada, bien para cumplir con programas de desarrollo o para responder a los estímulos procedentes de un medio ambiente cambiante. Una parte importante del genoma de las plantas está formado por genes que codifican factores de transcripción. Existen, en promedio, cerca de 2000 genes que codifican factores de transcripción en cada uno de los genomas de 165 especies distintas, según la "Plant Transcription Factor Database" (PlantTFDB v5.0, http://planttfdb.cbi.pku.edu.cn/). Los FT se unen a secuencias de ADN específicas, generalmente cortas, en las regiones promotoras (reguladoras) de los genes diana denominados elementos reguladores en *cis* (*cis*-reguladores). Las diferentes clases de FT poseen dominios característicos de unión al ADN que discriminan entre diversos elementos *cis*-reguladores y se unen a sus secuencias diana específicamente (Ciolkowski et al., 2008;Alves et al., 2014).

Los factores de transcripción WRKY constituyen una de las principales familias de FT en las plantas superiores, tanto por su extensión como en importancia biológica (Rushton et al., 2010). Aunque esta familia se ha expandido enormemente en plantas superiores, también se encuentran representantes en plantas primitivas (algas, helechos, musgos) y algunos protistas, estando ausente en procariotas y resto de eucariotas (Ulker and Somssich, 2004;Rinerson et al., 2015).

Las proteínas WRKY se caracterizan por poseer un dominio conservado, que abarca unos 60 aminoácidos, formado por un motivo heptapéptido de secuencia WRKYGQK, de la reciben el nombre, junto a un motivo tipo "dedos de zinc" característico de este grupo. Dependiendo del número de dominios WRKY y de las características del motivo "dedos de zinc", se clasifica a los miembros de la familia WRKY en tres grupos principales: grupos I, II y III (Figura 11A). El Grupo I, está formado por proteínas que contienen dos dominios WRKY situados en las posiciones N- y C- terminales (I-NT y I-CT). A su vez, en el Grupo II pueden distinguirse hasta cinco subgrupos (IIa, IIb, IIc, IId y IIe). Esta clasificación se debe a diferencias conservadas en la composición de los aminoácidos que forman el dominio (Eulgem et al., 2000;Rushton et al., 2010).

El dominio WRKY es el responsable de la unión con el ADN, formándose una estructura terciaria, estabilizada por el motivo "dedos de zinc", en la que el motivo WRKYGQK se une al ADN reconociendo específicamente al elemento regulador denominado W-box (Maeo et al., 2001;Llorca et al., 2014) (Figura 11B). La secuencia consenso mínima de la W-box es TTGACY (siendo Y=C/T), aunque las secuencias adyacentes también contribuyen a la afinidad de la unión ADN-proteína y probablemente marquen cierto grado de especificidad (Ciolkowski et al., 2008).

Otras proteínas reguladoras de la transcripción no se unen al ADN directamente, sino que forman complejos proteína-proteína con factores de transcripción para modular la respuesta transcripcional de la célula. Así funciona un nuevo grupo de proteínas

reguladoras en plantas, recientemente identificado, denominadas proteínas valina-glutamina o proteínas VQ (Cheng et al., 2012). Todos sus miembros comparten un dominio de aminoácidos conservado, de secuencia FxxhVQxhTG, involucrado en la unión al dominio WRKY C-terminal (I-CT) de los factores de transcripción WRKY del grupo I y al dominio único de los del grupo IIc. De esta forma, las proteínas VQ modulan la actividad transcripcional de muchos factores WRKY, ejerciendo una acción activadora o represora de la misma (Jing and Lin, 2015). Aunque las proteínas VQ pueden interaccionar físicamente con otros factores de transcripción (como PIF1, ABI5 y algunos MYB) y proteínas reguladoras de transducción de señales (MAPKs, CaM1), las VQ parecen haber evolucionado para actuar, principalmente, como cofactores de factores de transcripción tipo WRKY (Cheng et al., 2012;Li et al., 2014). De esta forma, los complejos WRKY-VQ están implicados en la regulación de las respuestas de la planta a estreses bióticos y abióticos (Luhua et al., 2008;Lai et al., 2011;Kim et al., 2013;Wang et al., 2014;Yan et al., 2018) y regulan varios procesos relacionados con el crecimiento y desarrollo (Cheng et al., 2012;Li et al., 2014;Lei et al., 2017) (Figura 12).

Aunque aquí hemos introducido brevemente tanto a los factores de transcripción WRKY, como a las proteínas VQ, ambos serán tratados en mayor profundidad en los capítulos 2 y 3.

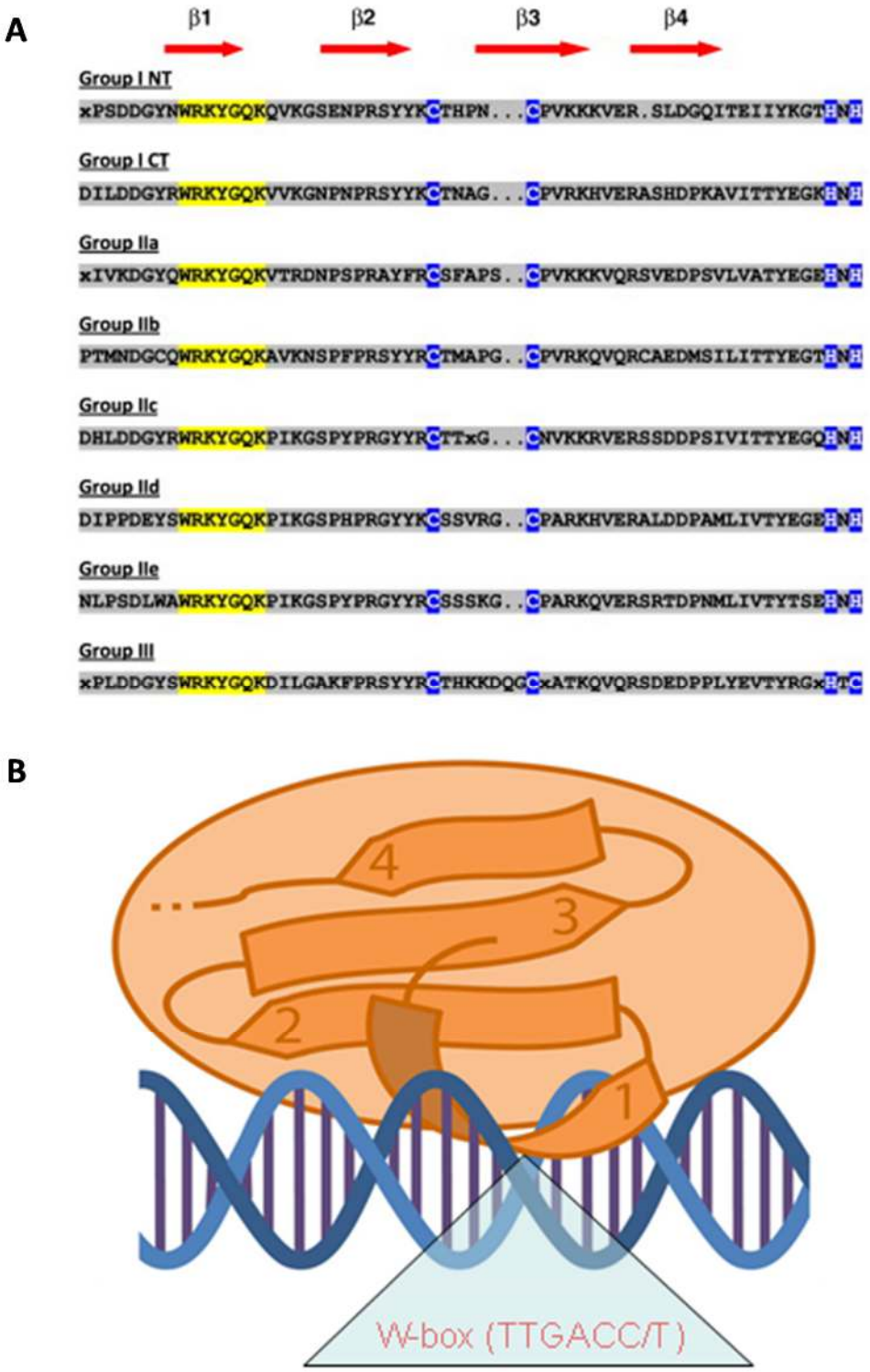


Figura 11. A, Estructura secundaria y secuencias consenso del dominio WRKY (63 aminoácidos) de cada uno de los diferentes grupos de FT tipo WRKY, señalando el motivo WRKY (amarillo) y los residuos de cisteína e histidina (azul) del motivo "dedos de zinc". **B**, Esquema de la interacción del dominio WRKY con el elemento *cis*-regulador W-box. β1 a β4, estructuras secundarias lamina-hoja plegada. El motivo WRKY, situado en la primera lámina β1, es el encargado del reconocimiento y unión a W-box, mientras que el motivo "dedos de zinc" estabiliza la estructura terciaria coordinando un átomo de zinc. Figura adaptada de Rushton et al. (2010) y Llorca et al. (2014).

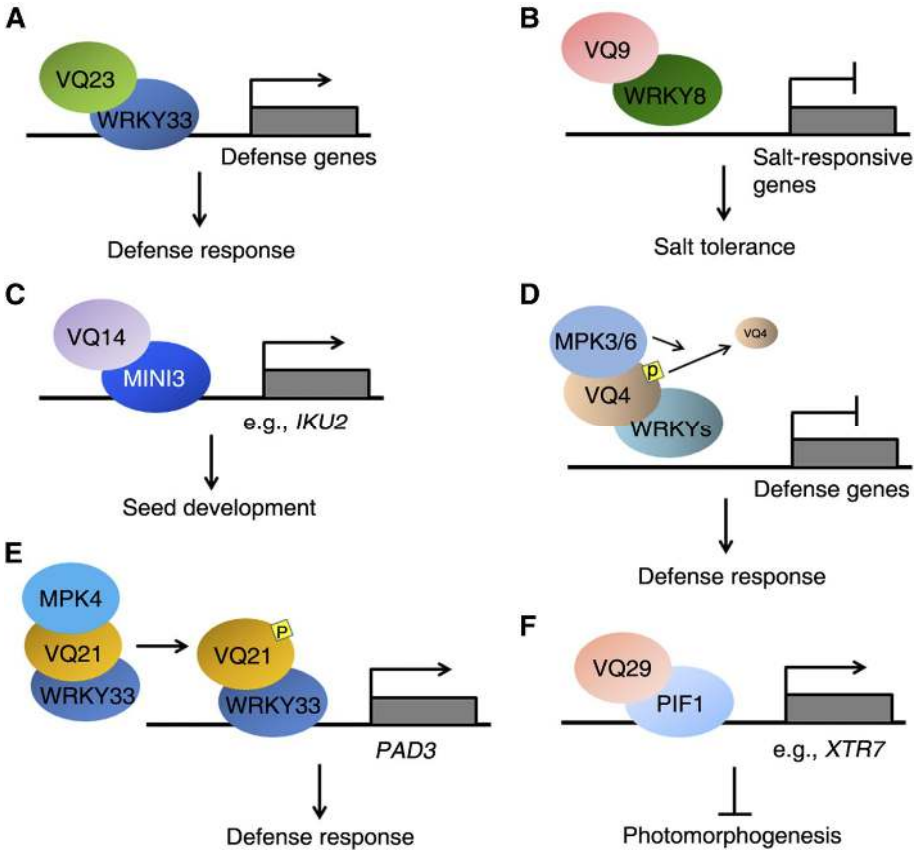


Figura 12. Modos de regulación transcripcional ejercida por proteínas VQ en *Arabidopsis thaliana*. **A**, VQ23 interactúa con WRKY33 y estimula su unión a la región promotora de genes que responden a la defensa. **B**, La interacción entre VQ9 y WRKY8 inhibe la unión de WRKY8 a las regiones promotoras de sus genes diana durante la respuesta de tolerancia a estrés salino. **C**, VQ14 y MINI3 interactúan y co-regulan el desarrollo de semillas. **D**, La interacción con VQ4 inhibe la unión de varios factores WRKY a genes diana relacionados con la respuesta de defensa de la planta. La fosforilación de VQ4, mediada por MPK3/6, induce su separación del complejo y su degradación, permitiendo la expresión de los genes regulados. **E**, Después de la infección por patógenos, MPK4 se activa y fosforila a VQ21 (MSK1), desencadenando la unión de VQ21 y WRKY33 a los promotores de genes de defensa, como *PAD3*, activando su expresión. **F**, VQ29 y PIF1 interactúan para regular la expresión de *XTR7* y promover la elongación celular regulada por luz. Figura tomada de Jing et al. (2015).

1.5. Objetivos

A partir de la información genómica más reciente se pretende identificar y caracterizar a todos los miembros de la familia de factores de transcripción WRKY en fresa, así como realizar un análisis exhaustivo de sus patrones de expresión espacio-temporal, fundamentado en los datos transcriptómicos disponibles para esta especie y en la bibliografía más actual.

Así mismo, se abordará la identificación y caracterización de la familia VQ, mucho menos conocida en fresa, mediante análisis de expresión génica de todos sus miembros en experimentos de respuesta a la infección con *C. acutatum* y tratamientos con las hormonas SA y MeJA, implicadas en la activación de rutas de defensa a patógenos. Se estudiarán las posibles interacciones proteína-proteína entre familias VQ y WRKY de fresa en la respuesta frente a *C. acutatum*.

Al mismo tiempo, se pretende esclarecer el papel biológico del factor de transcripción FaWRKY1 en la regulación de la respuesta de defensa del fruto de fresa frente a *C. acutatum*, mediante experimentos de silenciamiento y sobreexpresión transitoria del mismo.

Todo ello está orientado a proporcionar una información actualizada y valiosa del posible papel biológico que todos los miembros de estas familias de factores de transcripción puedan ejercer en fresa. Dicha información ayudará, sin duda, a seleccionar genes candidatos para futuros programas de mejora de la fresa tanto por métodos clásicos como mediante aproximaciones biotecnológicas novedosas y de vanguardia.

1.6. Bibliografía

Alkan, N., Friedlander, G., Ment, D., Prusky, D., and Fluhr, R. (2015). Simultaneous transcriptome analysis of Colletotrichum gloeosporioides and tomato fruit pathosystem reveals novel fungal pathogenicity and fruit defense strategies. *New Phytol* 205**,** 801-815.doi: 10.1111/nph.13087

Alves, M.S., Dadalto, S.P., Goncalves, A.B., De Souza, G.B., Barros, V.A., and Fietto, L.G. (2014). Transcription Factor Functional Protein-Protein Interactions in Plant Defense Responses. *Proteomes* 2**,** 85-106.doi: 10.3390/proteomes2010085

Amil-Ruiz, F. (2013). Molecular mechanisms of Strawberry Plant Defence against colletotrichum acutatum. PhD thesis, Universidad de Córdoba: Córdoba, Spain, 2013.

Amil-Ruiz, F., Blanco-Portales, R., Munoz-Blanco, J., and Caballero, J.L. (2011). The strawberry plant defense mechanism: a molecular review. *Plant Cell Physiol* 52**,** 1873-1903.doi: 10.1093/pcp/pcr136

Amil-Ruiz, F., Garrido-Gala, J., Gadea, J., Blanco-Portales, R., Munoz-Merida, A., Trelles, O., et al. (2016). Partial Activation of SA- and JA-Defensive Pathways in Strawberry upon Colletotrichum acutatum Interaction. *Front Plant Sci* 7**,** 1036.doi: 10.3389/fpls.2016.01036

Anciro, A., Mangandi, J., Verma, S., Peres, N., Whitaker, V.M., and Lee, S. (2018). FaRCg1: a quantitative trait locus conferring resistance to Colletotrichum crown rot caused by Colletotrichum gloeosporioides in octoploid strawberry. *Theor Appl Genet* 131**,** 2167-2177.doi: 10.1007/s00122-018-3145-z

Barbey, C.R., Lee, S., Verma, S., Bird, K.A., Yocca, A.E., Edger, P.P., et al. (2019). Disease Resistance Genetics and Genomics in Octoploid Strawberry. *G3 (Bethesda)* 9**,** 3315-3332.doi: 10.1534/g3.119.400597

Casado-Díaz, A., Encinas-Villarejo, S., Santos, B.D.L., Schilirò, E., Yubero-Serrano, E.-M., Amil-Ruíz, F., et al. (2006). Analysis of strawberry genes differentially expressed in response to Colletotrichum infection. *Physiologia Plantarum* 128**,** 633-650.doi: 10.1111/j.1399-3054.2006.00798.x

Chanderbali, A.S., Berger, B.A., Howarth, D.G., Soltis, D.E., and Soltis, P.S. (2017). Evolution of floral diversity: genomics, genes and gamma. *Philos Trans R Soc Lond B Biol Sci* 372.doi: 10.1098/rstb.2015.0509

Cheng, Y., Zhou, Y., Yang, Y., Chi, Y.J., Zhou, J., Chen, J.Y., et al. (2012). Structural and functional analysis of VQ motif-containing proteins in Arabidopsis as interacting proteins of WRKY transcription factors. *Plant Physiol* 159**,** 810-825.doi: 10.1104/pp.112.196816

Ciolkowski, I., Wanke, D., Birkenbihl, R.P., and Somssich, I.E. (2008). Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant Mol Biol* 68**,** 81-92.doi: 10.1007/s11103-008-9353-1

Cowger, C., and Brown, J.K.M. (2019). Durability of Quantitative Resistance in Crops: Greater Than We Know? *Annu Rev Phytopathol* 57**,** 253-277.doi: 10.1146/annurev-phyto-082718-100016

Crepet, W.L., Nixon, K.C., and Gandolfo, M.A. (2004). Fossil evidence and phylogeny: the age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceous deposits. *Am J Bot* 91**,** 1666-1682.doi: 10.3732/ajb.91.10.1666

Crouch, J., O'connell, R., Gan, P., Buiate, E., Torres, M., Beirn, L., et al. (2014). Genomics of plant-associated fungi: monocot pathogens.

Curry, K.J., Abril, M., Avant, J.B., and Smith, B.J. (2002). Strawberry Anthracnose: Histopathology of Colletotrichum acutatum and C. fragariae. *Phytopathology* 92**,** 1055-1063.doi: 10.1094/PHYTO.2002.92.10.1055

Darwish, O., Shahan, R., Liu, Z., Slovin, J.P., and Alkharouf, N.W. (2015). Re-annotation of the woodland strawberry (Fragaria vesca) genome. *BMC Genomics* 16**,** 29.doi: 10.1186/s12864-015-1221-1

De Queiroz, C.B., Correia, H.L.N., Santana, M.F., Batista, D.S., Vidigal, P.M.P., Brommonschenkel, S.H., et al. (2019). The repertoire of effector candidates in Colletotrichum lindemuthianum reveals important information about Colletotrichum genus lifestyle. *Appl Microbiol Biotechnol* 103**,** 2295-2309.doi: 10.1007/s00253-019-09639-9

Dean, R., Van Kan, J.A., Pretorius, Z.A., Hammond-Kosack, K.E., Di Pietro, A., Spanu, P.D., et al. (2012). The Top 10 fungal pathogens in molecular plant pathology. *Mol Plant Pathol* 13**,** 414-430.doi: 10.1111/j.1364-3703.2011.00783.x

Denoyes-Rothan, B., Guerin, G., Lerceteau-Kohler, E., and Risser, G. (2005). Inheritance of Resistance to Colletotrichum acutatum in Fragaria x ananassa. *Phytopathology* 95**,** 405-412.doi: 10.1094/PHYTO-95-0405

Edger, P.P., Mckain, M.R., Yocca, A.E., Knapp, S.J., Qiao, Q., and Zhang, T. (2020). Reply to: Revisiting the origin of octoploid strawberry. *Nat Genet* 52**,** 5-7.doi: 10.1038/s41588-019-0544-2

Edger, P.P., Poorten, T.J., Vanburen, R., Hardigan, M.A., Colle, M., Mckain, M.R., et al. (2019). Origin and evolution of the octoploid strawberry genome. *Nat Genet* 51**,** 541-547.doi: 10.1038/s41588-019-0356-4

Edger, P.P., Vanburen, R., Colle, M., Poorten, T.J., Wai, C.M., Niederhuth, C.E., et al. (2018). Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (Fragaria vesca) with chromosome-scale contiguity. *Gigascience* 7**,** 1-7.doi: 10.1093/gigascience/gix124

Eulgem, T., Rushton, P.J., Robatzek, S., and Somssich, I.E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci* 5**,** 199-206.doi: 10.1016/s1360-1385(00)01600-9

Feng, C., Wang, J., Harris, A.J., Folta, K.M., Zhao, M., and Kang, M. (2020). Tracing the diploid ancestry of the cultivated octoploid strawberry. *Mol Biol Evol*.doi: 10.1093/molbev/msaa238

Fierascu, R.C., Temocico, G., Fierascu, I., Ortan, A., and Babeanu, N.E. (2020). Fragaria Genus: Chemical Composition and Biological Activities. *Molecules* 25.doi: 10.3390/molecules25030498

Folta, K.M., and Barbey, C.R. (2019). The strawberry genome: a complicated past and promising future. *Hortic Res* 6**,** 97.doi: 10.1038/s41438-019-0181-z

Folta, K.M., and Davis, T.M. (2006). Strawberry Genes and Genomics. *Critical Reviews in Plant Sciences* 25**,** 399-415.doi: 10.1080/07352680600824831

Group, T.a.P. (2016). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* 181**,** 1-20.doi: 10.1111/boj.12385

Guidarelli, M., Carbone, F., Mourgues, F., Perrotta, G., Rosati, C., Bertolini, P., et al. (2011). Colletotrichum acutatum interactions with unripe and ripe strawberry fruits and differential responses at histological and transcriptional levels. *Plant Pathology* 60**,** 685-697.doi: 10.1111/j.1365-3059.2010.02423.x

Guidarelli, M., Zoli, L., Orlandini, A., Bertolini, P., and Baraldi, E. (2014). The mannose-binding lectin gene FaMBL1 is involved in the resistance of unripe strawberry fruits to Colletotrichum acutatum. *Mol Plant Pathol* 15**,** 832-840.doi: 10.1111/mpp.12143

Hardigan, M.A., Feldmann, M.J., Lorant, A., Bird, K.A., Famula, R., Acharya, C., et al. (2019). Genome Synteny Has Been Conserved Among the Octoploid Progenitors of Cultivated Strawberry Over Millions of Years of Evolution. *Front Plant Sci* 10**,** 1789.doi: 10.3389/fpls.2019.01789

He, C., Duan, K., Zhang, L., Zhang, L., Song, L., Yang, J., et al. (2019). Fast Quenching the Burst of Host Salicylic Acid Is Common in Early Strawberry/Colletotrichum fructicola Interaction. *Phytopathology* 109**,** 531-541.doi: 10.1094/PHYTO-02-18-0043-R

Hirakawa, H., Shirasawa, K., Kosugi, S., Tashiro, K., Nakayama, S., Yamada, M., et al. (2014). Dissection of the octoploid strawberry genome by deep sequencing of the genomes of Fragaria species. *DNA Res* 21**,** 169-181.doi: 10.1093/dnares/dst049

Horowitz, S., Freeman, S., and Sharon, A. (2002). Use of Green Fluorescent Protein-Transgenic Strains to Study Pathogenic and Nonpathogenic Lifestyles in Colletotrichum acutatum. *Phytopathology* 92**,** 743-749.doi: 10.1094/PHYTO.2002.92.7.743

Hummer, K.E., and Hancock, J. (2009). "Strawberry Genomics: Botanical History, Cultivation, Traditional Breeding, and New Technologies," in *Genetics and Genomics of Rosaceae,* eds. K.M. Folta & S.E. Gardiner. (New York, NY: Springer New York), 413-435.

Hummer, K.E., and Janick, J. (2009). "Rosaceae: Taxonomy, Economic Importance, Genomics," in *Genetics and Genomics of Rosaceae,* eds. K.M. Folta & S.E. Gardiner. (New York, NY: Springer New York), 1-17.

Irieda, H., Maeda, H., Akiyama, K., Hagiwara, A., Saitoh, H., Uemura, A., et al. (2014). Colletotrichum orbiculare Secretes Virulence Effectors to a Biotrophic Interface at the Primary Hyphal Neck via Exocytosis Coupled with SEC22-Mediated Traffic. *Plant Cell* 26**,** 2265-2281.doi: 10.1105/tpc.113.120600

Jacobs, R.L., Adhikari, T.B., Pattison, J., Yencho, G.C., Fernandez, G.E., and Louws, F.J. (2019). Inheritance of Resistance to Colletotrichum gloeosporioides and C. acutatum in Strawberry. *Phytopathology* 109**,** 428-435.doi: 10.1094/PHYTO-08-18-0283-R

Jayawardena, R.S. (2016). Notes on currently accepted species of Colletotrichum. *Mycosphere* 7**,** 1192-1260.doi: 10.5943/mycosphere/si/2c/9

Jiao, Y., Leebens-Mack, J., Ayyampalayam, S., Bowers, J.E., Mckain, M.R., Mcneal, J., et al. (2012). A genome triplication associated with early diversification of the core eudicots. *Genome Biol* 13**,** R3.doi: 10.1186/gb-2012-13-1-r3

Jing, Y., and Lin, R. (2015). The VQ Motif-Containing Protein Family of Plant-Specific Transcriptional Regulators. *Plant Physiol* 169**,** 371-378.doi: 10.1104/pp.15.00788

Jones, J.D., and Dangl, J.L. (2006). The plant immune system. *Nature* 444**,** 323-329.doi: 10.1038/nature05286

Jung, S., Lee, T., Cheng, C.H., Buble, K., Zheng, P., Yu, J., et al. (2019). 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. *Nucleic Acids Res* 47**,** D1137-D1145.doi: 10.1093/nar/gky1000

Kim, D.Y., Kwon, S.I., Choi, C., Lee, H., Ahn, I., Park, S.R., et al. (2013). Expression analysis of rice VQ genes in response to biotic and abiotic stresses. *Gene* 529**,** 208-214.doi: 10.1016/j.gene.2013.08.023

Lai, Z., Li, Y., Wang, F., Cheng, Y., Fan, B., Yu, J.Q., et al. (2011). Arabidopsis sigma factor binding proteins are activators of the WRKY33 transcription factor in plant defense. *Plant Cell* 23**,** 3824-3841.doi: 10.1105/tpc.111.090571

Lee, H.A., and Yeom, S.I. (2015). Plant NB-LRR proteins: tightly regulated sensors in a complex manner. *Brief Funct Genomics* 14**,** 233-242.doi: 10.1093/bfgp/elv012

Lei, R., Li, X., Ma, Z., Lv, Y., Hu, Y., and Yu, D. (2017). Arabidopsis WRKY2 and WRKY34 transcription factors interact with VQ20 protein to modulate pollen development and function. *Plant J* 91**,** 962-976.doi: 10.1111/tpj.13619

Lerceteau-Kohler, E., Guerin, G., and Denoyes-Rothan, B. (2005). Identification of SCAR markers linked to Rca 2 anthracnose resistance gene and their assessment in strawberry germ plasm. *Theor Appl Genet* 111**,** 862-870.doi: 10.1007/s00122-005-0008-1

Li, Y., Jing, Y., Li, J., Xu, G., and Lin, R. (2014). Arabidopsis VQ MOTIF-CONTAINING PROTEIN29 represses seedling deetiolation by interacting with PHYTOCHROME-INTERACTING FACTOR1. *Plant Physiol* 164**,** 2068-2080.doi: 10.1104/pp.113.234492

Li, Y., Pi, M., Gao, Q., Liu, Z., and Kang, C. (2019). Updated annotation of the wild strawberry Fragaria vesca V4 genome. *Hortic Res* 6**,** 61.doi: 10.1038/s41438-019-0142-6

Liang, X., Shang, S., Dong, Q., Wang, B., Zhang, R., Gleason, M.L., et al. (2018). Transcriptomic analysis reveals candidate genes regulating development and host interactions of Colletotrichum fructicola. *BMC Genomics* 19**,** 557.doi: 10.1186/s12864-018-4934-0

Liston, A., Cronn, R., and Ashman, T.L. (2014). Fragaria: a genus with deep historical roots and ripe for evolutionary and ecological insights. *Am J Bot* 101**,** 1686-1699.doi: 10.3732/ajb.1400140

Liston, A., Wei, N., Tennessen, J.A., Li, J., Dong, M., and Ashman, T.L. (2020). Revisiting the origin of octoploid strawberry. *Nat Genet* 52**,** 2-4.doi: 10.1038/s41588-019-0543-3

Llorca, C.M., Potschin, M., and Zentgraf, U. (2014). bZIPs and WRKYs: two large transcription factor families executing two different functional strategies. *Front Plant Sci* 5**,** 169.doi: 10.3389/fpls.2014.00169

Lo Iacono, G., Van Den Bosch, F., and Gilligan, C.A. (2013). Durable resistance to crop pathogens: an epidemiological framework to predict risk under uncertainty. *PLoS Comput Biol* 9**,** e1002870.doi: 10.1371/journal.pcbi.1002870

Luhua, S., Ciftci-Yilmaz, S., Harper, J., Cushman, J., and Mittler, R. (2008). Enhanced tolerance to oxidative stress in transgenic Arabidopsis plants expressing proteins of unknown function. *Plant Physiol* 148**,** 280-292.doi: 10.1104/pp.108.124875

Maas, J.L. (2004). "Strawberry Disease Management," in *Diseases of Fruits and Vegetables: Volume II,* ed. S.a.M.H. Naqvi. (Dordrecht: Springer Netherlands), 441-483.

Maeo, K., Hayashi, S., Kojima-Suzuki, H., Morikami, A., and Nakamura, K. (2001). Role of conserved residues of the WRKY domain in the DNA-binding of tobacco WRKY family proteins. *Biosci Biotechnol Biochem* 65**,** 2428-2436.doi: 10.1271/bbb.65.2428

Martin, R.R., and Tzanetakis, I.E. (2013). High Risk Strawberry Viruses by Region in the United States and Canada: Implications for Certification, Nurseries, and Fruit Production. *Plant Dis* 97**,** 1358-1362.doi: 10.1094/PDIS-09-12-0842-RE

Medina-Puche, L., Blanco-Portales, R., Molina-Hidalgo, F.J., Cumplido-Laso, G., Garcia-Caparros, N., Moyano-Canete, E., et al. (2016). Extensive transcriptomic studies on the roles played by abscisic acid and auxins in the development and ripening of strawberry fruits. *Funct Integr Genomics* 16**,** 671-692.doi: 10.1007/s10142-016-0510-3

Mengiste, T. (2012). Plant immunity to necrotrophs. *Annu Rev Phytopathol* 50**,** 267-294.doi: 10.1146/annurev-phyto-081211-172955

Mezzetti, B., Giampieri, F., Zhang, Y.-T., and Zhong, C.-F. (2018). Status of strawberry breeding programs and cultivation systems in Europe and the rest of the world. *Journal of Berry Research* 8**,** 205-221.doi: 10.3233/jbr-180314

Molina-Hidalgo, F.J., Franco, A.R., Villatoro, C., Medina-Puche, L., Mercado, J.A., Hidalgo, M.A., et al. (2013). The strawberry (Fragariaxananassa) fruit-specific rhamnogalacturonate lyase 1 (FaRGLyase1) gene encodes an enzyme involved in the degradation of cell-wall middle lamellae. *J Exp Bot* 64**,** 1471-1483.doi: 10.1093/jxb/ers386

Munch, S., Lingner, U., Floss, D.S., Ludwig, N., Sauer, N., and Deising, H.B. (2008). The hemibiotrophic lifestyle of Colletotrichum species. *J Plant Physiol* 165**,** 41-51.doi: 10.1016/j.jplph.2007.06.008

Nellist, C.F. (2018). "Disease Resistance in Polyploid Strawberry," in *The Genomes of Rosaceous Berries and Their Wild Relatives,* eds. T. Hytönen, J. Graham & R. Harrison. (Cham: Springer International Publishing), 79-94.

Njuguna, W., Liston, A., Cronn, R., Ashman, T.L., and Bassil, N. (2013). Insights into phylogeny, sex function and age of Fragaria based on whole chloroplast genome sequencing. *Mol Phylogenet Evol* 66**,** 17-29.doi: 10.1016/j.ympev.2012.08.026

O'connell, R.J., Thon, M.R., Hacquard, S., Amyotte, S.G., Kleemann, J., Torres, M.F., et al. (2012). Lifestyle transitions in plant pathogenic Colletotrichum fungi deciphered by genome and transcriptome analyses. *Nat Genet* 44**,** 1060-1065.doi: 10.1038/ng.2372

Paniagua, C., Santiago-Domenech, N., Kirby, A.R., Gunning, A.P., Morris, V.J., Quesada, M.A., et al. (2017). Structural changes in cell wall pectins during strawberry fruit development. *Plant Physiol Biochem* 118**,** 55-63.doi: 10.1016/j.plaphy.2017.06.001

Peres, N.A., Timmer, L.W., Adaskaveg, J.E., and Correll, J.C. (2005). Lifestyles of Colletotrichum acutatum. *Plant Dis* 89**,** 784-796.doi: 10.1094/PD-89-0784

Pieterse, C.M., Van Der Does, D., Zamioudis, C., Leon-Reyes, A., and Van Wees, S.C. (2012). Hormonal modulation of plant immunity. *Annu Rev Cell Dev Biol* 28**,** 489-521.doi: 10.1146/annurev-cellbio-092910-154055

Prusky, D., Mcevoy, J.L., Leverentz, B., and Conway, W.S. (2001). Local modulation of host pH by Colletotrichum species as a mechanism to increase virulence. *Mol Plant Microbe Interact* 14**,** 1105-1113.doi: 10.1094/MPMI.2001.14.9.1105

Rahman, T.A., Oirdi, M.E., Gonzalez-Lamothe, R., and Bouarab, K. (2012). Necrotrophic pathogens use the salicylic acid signaling pathway to promote disease development in tomato. *Mol Plant Microbe Interact* 25**,** 1584-1593.doi: 10.1094/MPMI-07-12-0187-R

Rinerson, C.I., Rabara, R.C., Tripathi, P., Shen, Q.J., and Rushton, P.J. (2015). The evolution of WRKY transcription factors. *BMC Plant Biol* 15**,** 66.doi: 10.1186/s12870-015-0456-y

Robert-Seilaniantz, A., Grant, M., and Jones, J.D. (2011). Hormone crosstalk in plant disease and defense: more than just jasmonate-salicylate antagonism. *Annu Rev Phytopathol* 49**,** 317-343.doi: 10.1146/annurev-phyto-073009-114447

Rushton, P.J., Somssich, I.E., Ringler, P., and Shen, Q.J. (2010). WRKY transcription factors. *Trends Plant Sci* 15**,** 247-258.doi: 10.1016/j.tplants.2010.02.006

Salinas, N., Verma, S., Peres, N., and Whitaker, V.M. (2019). FaRCa1: a major subgenome-specific locus conferring resistance to Colletotrichum acutatum in strawberry. *Theor Appl Genet* 132**,** 1109-1120.doi: 10.1007/s00122-018-3263-7

Shulaev, V., Sargent, D.J., Crowhurst, R.N., Mockler, T.C., Folkerts, O., Delcher, A.L., et al. (2011). The genome of woodland strawberry (Fragaria vesca). *Nat Genet* 43**,** 109-116.doi: 10.1038/ng.740

Sobczyk, M.K. (2018). "Phylogenetics of Fragaria, Rubus and Related Taxa," in *The Genomes of Rosaceous Berries and Their Wild Relatives,* eds. T. Hytönen, J. Graham & R. Harrison. (Cham: Springer International Publishing), 11-23.

Soltis, P.S., Folk, R.A., and Soltis, D.E. (2019). Darwin review: angiosperm phylogeny and evolutionary radiations. *Proceedings of the Royal Society B: Biological Sciences* 286**,** 20190099.doi: 10.1098/rspb.2019.0099

Takahara, H., Hacquard, S., Kombrink, A., Hughes, H.B., Halder, V., Robin, G.P., et al. (2016). Colletotrichum higginsianum extracellular LysM proteins play dual roles in appressorial function and suppression of chitin-triggered plant immunity. *New Phytol* 211**,** 1323-1337.doi: 10.1111/nph.13994

Tennessen, J.A., Govindarajulu, R., Ashman, T.L., and Liston, A. (2014). Evolutionary origins and dynamics of octoploid strawberry subgenomes revealed by dense targeted capture linkage maps. *Genome Biol Evol* 6**,** 3295-3313.doi: 10.1093/gbe/evu261

Tomas-Grau, R.H., Di Peto, P., Chalfoun, N.R., Grellet-Bournonville, C.F., Martos, G.G., Debes, M., et al. (2019). Colletotrichum acutatum M11 can suppress the defence response in strawberry plants. *Planta* 250**,** 1131-1145.doi: 10.1007/s00425-019-03203-5

Ulker, B., and Somssich, I.E. (2004). WRKY transcription factors: from DNA binding towards biological function. *Curr Opin Plant Biol* 7**,** 491-498.doi: 10.1016/j.pbi.2004.07.012

Van Eck, L., and Bradeen, J.M. (2018). "The NB-LRR Disease Resistance Genes of Fragaria and Rubus," in *The Genomes of Rosaceous Berries and Their Wild Relatives*.), 63-75.

Vergauwen, D., and De Smet, I. (2019). The Strawberry Tales: Size Matters. *Trends Plant Sci* 24**,** 1-3.doi: 10.1016/j.tplants.2018.10.007

Verma, S., Osorio, L.F., Lee, S., Bassil, N.V., and Whitaker, V.M. (2018). "Genome-Assisted Breeding in the Octoploid Strawberry," in *The Genomes of Rosaceous Berries and Their*

Wild Relatives, eds. T. Hytönen, J. Graham & R. Harrison. (Cham: Springer International Publishing), 161-184.

Wang, F., Zhang, F., Chen, M., Liu, Z., Zhang, Z., Fu, J., et al. (2017). Comparative Transcriptomics Reveals Differential Gene Expression Related to Colletotrichum gloeosporioides Resistance in the Octoploid Strawberry. *Front Plant Sci* 8**,** 779.doi: 10.3389/fpls.2017.00779

Wang, H., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R., et al. (2009). Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proc Natl Acad Sci U S A* 106**,** 3853-3858.doi: 10.1073/pnas.0813376106

Wang, X., Zhang, H., Sun, G., Jin, Y., and Qiu, L. (2014). Identification of active VQ motif-containing genes and the expression patterns under low nitrogen treatment in soybean. *Gene* 543**,** 237-243.doi: 10.1016/j.gene.2014.04.012

Wang, Y., Li, W., Chang, H., Zhou, J., Luo, Y., Zhang, K., et al. (2020). SRNAome and transcriptome analysis provide insight into strawberry fruit ripening. *Genomics* 112**,** 2369-2378.doi: 10.1016/j.ygeno.2020.01.008

Whitaker, V.M., Knapp, S.J., Hardigan, M.A., Edger, P.P., Slovin, J.P., Bassil, N.V., et al. (2020). A roadmap for research in octoploid strawberry. *Hortic Res* 7**,** 33.doi: 10.1038/s41438-020-0252-1

Xiang, Y., Huang, C.H., Hu, Y., Wen, J., Li, S., Yi, T., et al. (2017). Evolution of Rosaceae Fruit Types Based on Nuclear Phylogeny in the Context of Geological Times and Genome Duplication. *Mol Biol Evol* 34**,** 262-281.doi: 10.1093/molbev/msw242

Yan, C., Fan, M., Yang, M., Zhao, J., Zhang, W., Su, Y., et al. (2018). Injury Activates Ca(2+)/Calmodulin-Dependent Phosphorylation of JAV1-JAZ8-WRKY51 Complex for Jasmonate Biosynthesis. *Mol Cell* 70**,** 136-149 e137.doi: 10.1016/j.molcel.2018.03.013

Yang, J., Duan, G., Li, C., Liu, L., Han, G., Zhang, Y., et al. (2019). The Crosstalks Between Jasmonic Acid and Other Plant Hormone Signaling Highlight the Involvement of Jasmonic Acid as a Core Component in Plant Response to Biotic and Abiotic Stresses. *Front Plant Sci* 10**,** 1349.doi: 10.3389/fpls.2019.01349

Zhang, L., Huang, X., He, C., Zhang, Q.Y., Zou, X., Duan, K., et al. (2018). Novel Fungal Pathogenicity and Leaf Defense Strategies Are Revealed by Simultaneous Transcriptome Analysis of Colletotrichum fructicola and Strawberry Infected by This Fungus. *Front Plant Sci* 9**,** 434.doi: 10.3389/fpls.2018.00434

Zhang, L., Zhang, F., Melotto, M., Yao, J., and He, S.Y. (2017). Jasmonate signaling and manipulation by pathogens and insects. *J Exp Bot* 68**,** 1371-1385.doi: 10.1093/jxb/erw478

Zhang, Q.Y., Zhang, L.Q., Song, L.L., Duan, K., Li, N., Wang, Y.X., et al. (2016). The different interactions of Colletotrichum gloeosporioides with two strawberry varieties and the involvement of salicylic acid. *Hortic Res* 3**,** 16007.doi: 10.1038/hortres.2016.7

Capítulo 2.

A comprehensive study of the WRKY transcription factor family in strawberry

Capítulo 2. A comprehensive study of the WRKY transcription factor family in strawberry

2.1. Abstract

WRKY transcription factors play critical roles in plant growth and development or stress responses. Using up-to date genomic data, in this work has been identified a total of of 64 and 257 WRKY transcription factors in the diploid *Fragaria vesca* and the cultivated allo-octoploid *Fragaria x ananassa* cv. Camarosa. The completeness of the new genomes and annotations has enabled us to perform a more detailed evolutionary and functional study of the strawberry WRKY family members, particularly in the case of the cultivated hybrid, in which homoelogous and paralogous *FaWRKY* genes have been characterized. Analysis of the available expression profiles have revealed that many strawberry *WRKY* genes showed preferential or tissue-specific expression. Furthermore, significant differential expression of several FaWRKY genes has been clearly detected in fruit receptacles and achenes during the ripening process and pathogen challenged, supporting a precise functional role of these strawberry genes in such processes. Our results provide a deeper and more comprehensive knowledge of the WRKY gene family in strawberry.

2.2. Introduction

The WRKY superfamily of transcription factors (WRKY TFs) (Eulgem et al., 2000) has an early origin in primitive eukaryotes, with later expansion and evolution in the green lineage driven by extensive tandem and segmental duplication events (Zhang and Wang, 2005;Bakshi and Oelmuller, 2014;Mohanta et al., 2016), thus becoming one of the largest TF families in higher plants (Rushton et al., 2010). The WRKY TFs are involved in the regulation of various physiological and developmental processes, such as senescence (Besseau et al., 2012;Chen et al., 2017b;Doll et al., 2019), stem elongation and seed development (Zhang et al., 2011), or flowering (Zhang et al., 2018a). They are also key players in responses to abiotic stresses (Chen et al., 2012) as well as to biotic stresses, wherein they seem to play major roles in plant immunity (Rushton et al., 2010;Alves et al., 2014;Tsuda and Somssich, 2015). In *Arabidopsis*, several WRKY TFs, such as WRKY3, -4, -6, -29, -33, -52 and -70 have been identified as positive regulators of defense responses against pathogenic fungi and bacteria whereas several others, such as WRKY7, -11, -17, -18, -23, -25, -27, -38, -40, -41, -48, -53, -58, -60 and -62 have been found to play negative regulatory roles (Pandey and Somssich, 2009). However, many WRKY TFs exhibit a dual activity in plant defense, depending on the type of response to the pathogen. Thus, *Arabidopsis* WRKY70 plays an important role balancing the SA- and JA- dependent defense response pathways, inducing SA-responsive PR genes to enhance the resistance to biotrophic pathogens, at

the time that repress the expression of JA-responsive genes, compromising resistance to necrotrophs (Li et al., 2004;Li et al., 2006;Li et al., 2017). On the other hand, overexpression of *AtWRKY33*, a negative regulator of SA-pathway responses, increases resistance to fungal necrotrophic pathogens, but enhances susceptibility to the bacterial pathogen *P. syringae* (Zheng et al., 2006).

WRKY TFs regulate their target's expression by binding to a specific cis-element known as W-box, with the minimal consensus sequence TTGACY (where Y=C/T), although adjacent sequences are also involved in the binding site preferences (Ciolkowski et al., 2008). DNA binding by WRKY proteins is mediated by a highly conserved DNA binding domain, named the WRKY domain (WD). It is about 60 amino acids long and contains a N-terminal core motif, formed by the almost invariant heptapeptide WRKYGQK, and a distinctive C-terminal zinc-finger (Znf), both required for the DNA binding activity (Maeo et al., 2001). The amino acids constituting the WRKY core motif are essential for the DNA-binding activity and recognition specificity, thus amino acid substitutions in this motif can affect the binding to the W-box (Ciolkowski et al., 2008;Cheng et al., 2019;Singh et al., 2019). Variations in this conserved sequence have been found in WRKY proteins from different plant species (Mohanta et al., 2016), with functional and binding activity characterization in some cases. WRKYs harbouring divergent core motifs can bind to novel sequences that deviate from the consensus W-box. For example, NtWRKY12 contains a WRKYGKK core motif, interacting with the WK-box (TTTTCCAC) but not to the consensus W-box (van Verk et al., 2008). On the other hand, modified motifs can be also unable to bind to the W-box sequence, as GmWRKY167, containing the WRKYEDK core motif (Yang et al., 2017). However, it is unclear if the binding specificity depends exclusively on the WRKY core motif. OsWRKY7, which contains the WRKYGKK sequence, binds to the W-box but not with the WK-box (Chen et al., 2019). In addition, AtWRKY70, which harbour the conserved WRKYGQK core motif, can bind to consensus W-box, as well as the novel WT-box (YGACTTTT) (Machens et al., 2014).

The initial WRKY classification in three groups was based on both the number of WDs and the pattern of the Znf domain (Eulgem et al., 2000). Group I WRKY proteins are the only ones that have two WDs (I-NT, I-CT), with a C2H2 Znf pattern ($CX_{4-5}CX_{22-23}HXH$) shared with group II WRKYs. Group II was further divided into subgroups IIa, IIb, IIc, IId and IIe based on the differences in their amino acid sequences from their WD. Group III WRKY proteins contain a WD with a C2HC Znf pattern ($CX_7CX_{23}HXC$). This classification has been maintained over time, but the increasing availability of complete plant genomes has allowed a more detailed examination of its evolution in the plant lineage. Subsequent phylogenetic analyses proposed that an ancestral I-CT gene underwent domain duplication, resulting in plant WRKY Group I. Then, the other WRKY groups evolved from Group I genes lacking the I-NT domain (Zhang and Wang, 2005). These authors also showed that Group II is not monophyletic, so the WRKY family in higher plants could be classified into five groups. No changes were applied for Groups I and III, while the subgroups of Group II were further divided into Group IIc, Group IIa+IIb and Group IId+IIe.

Later studies have proposed a new phylogenetic classification into four major groups (Group I+IIc, Group IIa+IIb, Group IId+IIe and Group III), as well as two different hypotheses for the evolution of WRKY genes, the “Group I Hypothesis” and the “IIa + b Separate Hypothesis” (Rinerson et al., 2015). The “Group I Hypothesis” suggests that all WRKY genes in higher plants have evolved from an ancestral Group I gene, whereas the “IIa + b Separate Hypothesis” propose that Group IIa and IIb WRKY genes have evolved from a single domain algal WRKY gene separated from the lineage of Group I. More recently, marked phylogenetic differences have been observed between monocot and dicot WRKY TFs, and it has been proposed that the clustering system should be specific for monocot or dicot plant lineages, with up to six clusters for monocot WRKY TFs, Groups I to VI, and three clusters for dicot WRKY TFs, Groups I, II (sub-groups IIa, IIb, IIc) and III (Mohanta et al., 2016).

The WRKY proteins can also include additional domains. Notoriously, the R protein-WRKY family contains several typical domains of R-proteins, such as toll interleukin 1 receptor (TIR), leucine-rich repeat (LRR) and NB-ARC, which are associated with WDs from Group I, Group II and Group III members. Such chimeric genes have been found in multiple plant genomes, however they are not widespread in plants. Instead, R protein-WRKY genes appear to have evolved on multiple independent occasions as a result of particular genomic rearrangements within specific plant lineages (Rinerson et al., 2015). For instance, AtWRKY52 (RRS1) is a functionally characterized R-protein-WRKY involved in resistance against bacterial and fungal pathogens through pathogen-effector recognition and in association with the disease resistance protein RPS4 (Deslandes et al., 2003;Narusaka et al., 2009;Guo et al., 2020).

The genus *Fragaria*, (*Rosaceae*) comprises about 24 species worldwide, with different geographical distribution and ploidy levels, from diploid to decaploid (DiMeglio et al., 2014). The modern cultivated octoploid strawberry (*Fragaria x ananassa* Duch.) is presumably the most economically important soft berry, with a world production of more than 8.33 Mt and 372,361 ha harvested (FAOSTAT, 2018). In recent years, many efforts have been made to unravel the genetic background of this species, to be used as a molecular breeding tool to identify traits and associated genes of interest for the genetic improvement of this valuable crop.

The genome of *Fragaria vesca* (2n=14) was sequenced and published for the first time in 2011 (Shulaev et al., 2011) and proposed as a gateway to functional studies of genes within the *Rosaceae*, particularly for the cultivated strawberry. The first assembly versions were improved and reannotated (Tennessen et al., 2014;Darwish et al., 2015;Li et al., 2018). Recently, it has been sequenced *de novo* using third-generation PacBio Single Molecule, Real-Time (SMRT) sequencing technology, and an overall improvement over previous versions has been achieved (Edger et al., 2018). Moreover, gene models and genome annotation have been recently updated (Li et al., 2019).

Fragaria x ananassa is an allo-octoploid hibrid (2n=8x=56) originated, around 300 years ago, from interspecific crosses of the also octoploids *F. virginiana* and *F. chiloensis*. Recently, the genome of *Fragaria x ananassa* cv. Camarosa has been

completely sequenced and annotated, revealing its diploid progenitor species: *F. vesca* (subsp. *Bracheata*), *F. iinumae, F. nipponica and F. viridis* (Edger et al., 2019), although the contribution of the last two species remains in debate (Edger et al., 2020;Feng et al., 2020;Liston et al., 2020). Remarkably, this study has found that the subgenome contributed by *F. vesca* is dominant and has replaced large portions of the submissive ones through homoeologous exchanges.

The emergence of all these new data represents a valuable opportunity to conduct new and more comprehensive evolutionary analyses in strawberry, as well as greatly facilitate functional studies to further unravel the roles of members of the WRKY family in regulating the physiology of the strawberry and in responses to stress, particularly in key aspects such as fruit ripening and biotic stress. This is especially relevant for the cultivated strawberry, in which many previous studies were carried out with limited availability of genetic data on the structure and nucleotide sequence of the octoploid genome, but taking advantage of the synteny and high sequence identity with *F. vesca* to use its reference genome as “anchor” between both species (Amil-Ruiz et al., 2013;Amil-Ruiz et al., 2016;Medina-Puche et al., 2016;Sanchez-Sevilla et al., 2017). The present study updates and expands our knowledge of the members of diploid and octoploid strawberry WRKY TF family, their evolutionary history and their potentially roles in specific strawberry tissues and important biological processes such as fruit ripening and defense responses against pathogens.

2.3. Materials and Methods

2.3.1. Identification of WRKY Family members

The sequences of *Fragaria vesca* (Fv; Genome Assembly v4.0.a1 & Annotation v4.0.a2) and *Fragaria* x ananassa cv. Camarosa (Fa; Genome Assembly v1.0 & Annotation v1.0.a1) were retrieved from The Genome Database for Rosaceae (GDR) website (https://www.rosaceae.org/) (Jung et al., 2019). The Hidden Markov Model of the WD (PF03106) was downloaded from the Pfam database (El-Gebali et al., 2019) and used as query in HMMER3 search, performed in the freeware tool UGENE v1.21 with default settings (Okonechnikov et al., 2012). The candidate sequences were further confirmed to include the WD, as well as additional protein domains, using the Conserved Domain Database (CDD) (Marchler-Bauer et al., 2017). WRKY sequences from *Arabidopsis thaliana* (At) (Eulgem et al., 2000), *Vitis vinifera* (Vv) (Guo et al., 2014) and several species were retrieved from the Plant Transcription Factor Database (PlantTFDB) (Jin et al., 2017). Chromosome maps of the strawberry WRKY genes were drawn with MapChart v2.32 (Voorrips, 2002). Protein locations were predicted using LOCALIZER (Sperschneider et al., 2017). General sequence handling and WRKY protein properties calculation were performed using TBTools v1.055 (Chen et al., 2020) and the Freiburg Galaxy server (https://usegalaxy.eu/) (Afgan et al., 2016).

2.3.2. Phylogenomics

Full WRKY protein sequences were aligned by MUSCLE to generate unrooted phylogenetic trees by the Neighbor-Joining (N-J) method, in MEGA 7.0 (Kumar et al., 2016). The resulting protein trees were annotated with iTOL (Letunic and Bork, 2019). Compared synteny among species and strawberry WRKY gene duplications were studied in the CoGE web-platform using LAST to find gene homologies and SynMap2 or SynMap3D to find collinear blocks shared by two or three species, respectively (Table 1) (Lyons and Freeling, 2008;Haug-Baltzell et al., 2017). The non-synonymous (Kn) and synonymous (Ks) substitution rates between pairs of syntenic genes were calculated by codeml, implemented in SynMap, or in PAL2NAL (http://www.bork.embl.de/pal2nal/) (Suyama et al., 2006) for duplicated genes lacking syntenic conservation. The non-synonymous (Kn) and synonymous (Ks) substitutions were used to calculate the Kn/Ks ratios (ω) between paralogous strawberry WRKYs and thus estimate the selection pressure. Values of $\omega>1$ or $\omega<1$ indicate positive or purifying (negative) selection, respectively, while $\omega=1$ means neutral (absence of) evolution (Yang and Nielsen, 2002). Very low substitution values ($Ks < 0.01$) were considered as virtual absence of nucleotide mutation and thus not accounted to calculate ω, because it may result in inaccurate estimates (Villanueva-Canas et al., 2013;De La Torre et al., 2017).

eggNOG-Mapper v2 and the eggNOG 5.0 Database (Huerta-Cepas et al., 2017;Huerta-Cepas et al., 2019) were used to classify the WRKY orthologs in strawberry and other species, including *Amborella trichopoda, Juglans regia, Malus domestica* and *Glycine max*. GO terms with experimental evidence were acquired from the eggNOG results and plotted using WEGO 2.0 (Ye et al., 2018).

2.3.3. Expression analyses of diploid and octoploid strawberry *WRKYs*

Expression data of *FvWRKY genes* were taken from a previously published RNA-seq expression analysis in several *F. vesca* tissues and developmental stages (Li et al., 2019). Transcripts per Million (TPM) values were log10-transformed and depicted using the heatmap function of TBTools.

The expression patterns of the *FaWRKY* genes were obtained through a complete reanalysis of several *Fragaria x ananassa* RNAseq datasets, including strawberry plant and fruit tissues (Sanchez-Sevilla et al., 2017); and strawberry leaves infected by *C. fructicola* (Zhang et al., 2018b). Therefore, raw reads from the sequencing platform were processed to retain only high-quality sequences to be subsequently used for the mapping (Cutadapt v1.9, BBDuk v35.43). Sequencing adapters were first clipped from each library, and low-quality bases were trimmed. A Phred quality score of 24 was selected as threshold and reads with length less than 30nt were filtered out. Reads quality assessment was carried out using FastQC software (v0.11.8) to evaluate the effect of every step of this process. All subsequent analyses were conducted using these high-quality datasets. The remaining ribosomal RNA was detected by SortMeRNA software (v2.1) (Kopylova et al., 2012). Thus, adaptors clipped reads were

mapped to SortMeRNA prepackaged databases id98 (from Silva v119, and Rfam) with default parameters. A two-pass mode mapping was carried out by STAR (v2.7.3.a) (Dobin et al., 2013) with parameter "--quantMode GeneCounts" in the second pass to extract raw counts per annotated gene ID according to each particular library strandness. The obtained expression matrix was then supplied to the R library Deseq2 (v1.28.1) for a differential expression analysis (padj < 0.01 and absolute value of log2 fold change > 1). In addition, hierarchical clustering and heatmap analysis were done for those genes of interest. Quantification values were z-scored prior to the clustering analysis by Pearson correlation with method "complete" (hclust function in stats package from R). Additional figures were drawn using TBtools.

Table 1. Synteny analyses performed in CoGe and persistent links to results.

Analysis	Tool	CoGe link
F. vesca, *A. thaliana* and *V. vinifera*	SynMap3D	https://genomevolution.org/r/17qjv
A. thaliana and *F. vesca*	SynMap2	https://genomevolution.org/r/17jqo
V. vinifera and *F. vesca*	SynMap2	https://genomevolution.org/r/17jqs
F. vesca self-synteny	SynMap2	https://genomevolution.org/r/1ceiw
F. vesca and *F. x ananassa* (Camarosa) synteny and fractionation	SynMap2	https://genomevolution.org/r/18c2f
F. x ananassa (Camarosa) self-synteny (homoeologs and paralogs)	SynMap2	https://genomevolution.org/r/16dnh

2.4. Results and discussion

2.4.1. WRKY members of *F. vesca* and *F. ananassa*

The FvWRKY family has been described before using an earlier annotation (Wei et al., 2016;Zhou et al., 2016). Here, we have used the latest Fv genome and annotation versions available, which combines a higher quality reference genome (Edger et al., 2018) and high-fidelity gene models with RNA-seq support using expression data from different *F. vesca* accessions, tissues and fruit developmental stages (Li et al., 2019). A comprehensive list of the FvWRKY members across different Fv genome annotations is also provided (Supplementary Table S1). From the initial set of WRKY candidates identified by HMMER, a total of 64 *FvWRKY* coding genes and their splicing forms were confirmed in CDD (Supplementary Table S2), then named according with their chromosomal locations (Table 2). Some relevant properties of the FvWRKY proteins and their predicted subcellular location are also listed in Table 3. GRAVY (grand average of hydropathy) values are below 0, indicating that FvWRKYs are hydrophilic and more likely globular-shaped, while LOCALIZER predicted protein location into the nucleus for most FvWRKYs.

FvWRKY proteins were classified into groups I, II and III according to their WDs (Eulgem et al., 2000) (Table 2). Core WRKY motif modifications were detected in FvWRKY3 and -8 (WRKYGKK), FvWRKY21 (WKKYGQK), FvWRKY35 (WTKYDQR) and FvWRKY55 (WREYDQR). A more drastic modification is found in FvWRKY37, in which the core motif is found truncated and reduced to WRK. Moreover, differences in some *FvWRKYs* splicing forms (SF) were observed, affecting the nature of the encoded proteins. Thus, one SF from *FvWRKY20, -26* and *-43* (Group I) encode for WRKY proteins which have lost their I-NT WDs, while some SF from several *FvWRKYs* encode proteins containing incomplete WDs. Additionally, one SF from *FvWRKY54* and *-62* have lost their WDs completely. The regulation of all these alternative transcripts and whether they are efficiently translated into functional proteins remains to be studied.

Additional motifs were also found in some of the FvWRKY proteins (Figure 1). These include a Plant zinc cluster domain in FvWRKY2, -7, -9, -10 and -22. TIR, NB-ARC and leucine-rich repeat motifs (LRR) harbored by the FvRWRKY subfamily (Rinerson et al., 2015), consisting of FvWRKY35, FvWRKY55, FvWRKY61 and FvWRKY62. Besides, FvWRKY35 and -55 exhibit an additional WRKY-like domain, lacking the core motif but retaining the Znf portion. Additionally, the lost of the WD in the predicted *FvWRKY62.t6* splicing form would turn the translated product into a TIR-NBS-LRR protein.

Table 2. Fragaria vesca WRKY genes and proteins

Name	Gene id	Chr*	Group	Splicing Forms		WD modifications
				Number	Comments**	
FvWRKY1	FvH4_1g00960	Fvb1	I	2		
FvWRKY2	FvH4_1g16480	Fvb1	IId	2	incomplete WD in SF t2	
FvWRKY3	FvH4_1g22820	Fvb1	IIc	3		WRKYGKK
FvWRKY4	FvH4_1g23082	Fvb1	IIc	7	incomplete WD in SF t4 and t7	
FvWRKY5	FvH4_1g26200	Fvb1	IIb	2		
FvWRKY6	FvH4_1g26980	Fvb1	IIc	1		
FvWRKY7	FvH4_2g22300	Fvb2	IId	4		
FvWRKY8	FvH4_2g31400	Fvb2	IIc	2	incomplete WD in SF t2	WRKYGKK
FvWRKY9	FvH4_2g33920	Fvb2	IId	4	No WD in SF t2 and t4	
FvWRKY10	FvH4_2g36730	Fvb2	IId	2	incomplete WD in SF t2	
FvWRKY11	FvH4_2g41060	Fvb2	IIa	3		
FvWRKY12	FvH4_2g41070	Fvb2	IIa	3		
FvWRKY13	FvH4_3g01700	Fvb3	IIb	2		
FvWRKY14	FvH4_3g06200	Fvb3	IIc	1		
FvWRKY15	FvH4_3g07360	Fvb3	IIe	2		
FvWRKY16	FvH4_3g11140	Fvb3	IIe	1		
FvWRKY17	FvH4_3g23150	Fvb3	IIb	1		
FvWRKY18	FvH4_3g24010	Fvb3	I	2		
FvWRKY19	FvH4_3g39850	Fvb3	I	5		
FvWRKY20	FvH4_3g41430	Fvb3	I	9	SF t8 protein has only I-CT WD	
FvWRKY21	FvH4_3g45810	Fvb3	IIc	1		WKKYGQK
FvWRKY22	FvH4_4g06830	Fvb4	IId	2	incomplete WD in SF t2	
FvWRKY23	FvH4_4g20230	Fvb4	IIc	1		
FvWRKY24	FvH4_4g23480	Fvb4	IIc	1		
FvWRKY25	FvH4_4g30360	Fvb4	IIc	5		
FvWRKY26	FvH4_4g30640	Fvb4	I	4	SF t2 protein has only I-CT WD	
FvWRKY27	FvH4_4g31790	Fvb4	IIb	3		
FvWRKY28	FvH4_5g03210	Fvb5	IIe	2		
FvWRKY29	FvH4_5g04370	Fvb5	III	1		
FvWRKY30	FvH4_5g08610	Fvb5	IIc	1		
FvWRKY31	FvH4_5g15340	Fvb5	IIb	1		

Table 2 (continued)

Name	Gene id	Chr*	Group	Splicing Forms		WD modifications
				Number	Comments**	
FvWRKY32	FvH4_5g29820	Fvb5	IIb	3		
FvWRKY33	FvH4_5g39060	Fvb5	IIe	2		
FvWRKY34	FvH4_6g00500	Fvb6	I	1		
FvWRKY35	FvH4_6g01690	Fvb6	III	2		WTKYDQR
FvWRKY36	FvH4_6g02660	Fvb6	IId	2	incomplete WD in SF t2	
FvWRKY37	FvH4_6g09650	Fvb6	III	1		WRK
FvWRKY38	FvH4_6g09710	Fvb6	III	2		
FvWRKY39	FvH4_6g09730	Fvb6	III	1		
FvWRKY40	FvH4_6g09740	Fvb6	III	6		
FvWRKY41	FvH4_6g09750	Fvb6	III	1		
FvWRKY42	FvH4_6g09760	Fvb6	III	1		
FvWRKY43	FvH4_6g10510	Fvb6	I	8	SF t3 protein has only I-CT WD	
FvWRKY44	FvH4_6g13470	Fvb6	IIc	1		
FvWRKY45	FvH4_6g15050	Fvb6	I	9		
FvWRKY46	FvH4_6g27740	Fvb6	IIe	2		
FvWRKY47	FvH4_6g28650	Fvb6	IIa	3		
FvWRKY48	FvH4_6g38370	Fvb6	IIc	1		
FvWRKY49	FvH4_6g41210	Fvb6	I	1		
FvWRKY50	FvH4_6g42870	Fvb6	IIb	3		
FvWRKY51	FvH4_6g47800	Fvb6	I	3		
FvWRKY52	FvH4_6g49140	Fvb6	IIb	2		
FvWRKY53	FvH4_6g53770	Fvb6	IIc	1		
FvWRKY54	FvH4_7g10440	Fvb7	IIc	3	No WD in SF t2	
FvWRKY55	FvH4_7g11550	Fvb7	III	1		WREYDQR
FvWRKY56	FvH4_7g15350	Fvb7	IIe	1		
FvWRKY57	FvH4_7g15430	Fvb7	IIc	1		
FvWRKY58	FvH4_7g16150	Fvb7	III	1		
FvWRKY59	FvH4_7g26020	Fvb7	III	1		
FvWRKY60	FvH4_7g26030	Fvb7	III	1		
FvWRKY61	FvH4_7g26071	Fvb7	III	1		
FvWRKY62	FvH4_7g28550	Fvb7	III	12	No WD in SF t6	
FvWRKY63	FvH4_7g30460	Fvb7	IIe	1		
FvWRKY64	FvH4_7g31050	Fvb7	III	1		

**Chr: chromosome*

***WD: WRKY domain; SF: splicing form*

Table 3. FvWRKY protein properties.

Protein	MW (Da)	Isoelectric point	GRAVY	Length (aa)	Instability index	Cellular location
FvWRKY1	50995.380	6.684	-0.733	467	58.936	Nucleus
FvWRKY2	34709.792	9.612	-0.584	319	52.081	Nucleus
FvWRKY3	22969.113	6.153	-0.830	205	42.612	-
FvWRKY4	27804.891	9.082	-0.807	243	54.189	Nucleus
FvWRKY5	46697.077	6.786	-0.574	425	51.925	-
FvWRKY6	35377.207	6.662	-0.747	317	66.553	Nucleus
FvWRKY7	38820.560	9.710	-0.749	347	56.734	Nucleus
FvWRKY8	17776.157	5.134	-1.139	155	49.850	-
FvWRKY9	35641.876	9.639	-0.555	329	56.045	Nucleus
FvWRKY10	37059.501	9.511	-0.570	342	51.885	Nucleus
FvWRKY11	42174.802	6.273	-0.649	381	54.188	Nucleus
FvWRKY12	36683.872	8.889	-0.646	332	47.005	Nucleus
FvWRKY13	61543.322	7.147	-0.742	571	49.185	-
FvWRKY14	39181.559	6.999	-0.943	355	58.416	Nucleus
FvWRKY15	30429.242	5.187	-0.789	277	65.152	Nucleus
FvWRKY16	53620.277	5.919	-0.740	497	51.895	Nucleus
FvWRKY17	59559.813	6.220	-0.590	547	40.266	Nucleus
FvWRKY18	62376.133	6.729	-0.847	573	57.698	Nucleus
FvWRKY19	68675.600	7.675	-1.046	627	49.750	Nucleus
FvWRKY20	79159.045	5.866	-0.788	734	51.545	Nucleus
FvWRKY21	15368.180	5.645	-1.335	137	34.369	-
FvWRKY22	32897.847	9.968	-0.674	301	53.717	Nucleus
FvWRKY23	41486.801	5.685	-0.930	378	58.879	Nucleus
FvWRKY24	21670.994	9.451	-0.878	190	41.592	Nucleus
FvWRKY25	34188.982	6.454	-0.915	313	65.874	Nucleus
FvWRKY26	54764.485	7.693	-1.029	499	57.489	Nucleus
FvWRKY27	67726.095	6.672	-0.821	615	46.461	Nucleus
FvWRKY28	35224.823	5.142	-0.777	309	61.021	Nucleus
FvWRKY29	40087.948	4.958	-0.755	355	63.992	Nucleus
FvWRKY30	18182.046	9.133	-0.872	159	50.069	Chloroplast, Nucleus
FvWRKY31	72804.217	6.953	-0.834	672	55.831	Nucleus
FvWRKY32	58092.742	5.019	-0.986	521	60.355	-
FvWRKY33	32380.267	4.917	-1.126	282	69.815	Nucleus
FvWRKY34	80231.150	5.875	-0.753	727	50.051	Nucleus
FvWRKY35	167206.175	6.103	-0.326	1476	45.783	Nucleus

Table 3 (continued)

Protein	MW (Da)	Isoelectric point	GRAVY	Length (aa)	Instability index	Cellular location
FvWRKY36	35767.547	9.576	-0.641	319	55.297	Nucleus
FvWRKY37	34798.233	5.232	-0.490	312	52.885	Nucleus
FvWRKY38	33543.344	7.168	-0.564	297	51.692	Nucleus
FvWRKY39	37898.521	5.262	-0.683	335	47.433	Nucleus
FvWRKY40	25954.892	9.316	-0.659	231	49.837	Nucleus
FvWRKY41	38756.337	5.413	-0.729	344	48.904	Nucleus
FvWRKY42	38214.851	5.637	-0.761	339	46.038	-
FvWRKY43	57462.961	7.633	-1.016	517	62.504	Nucleus
FvWRKY44	33307.614	5.438	-0.795	296	58.729	Nucleus
FvWRKY45	52189.199	9.143	-0.889	478	49.796	Nucleus
FvWRKY46	29902.823	5.120	-0.797	268	53.138	-
FvWRKY47	36079.079	8.799	-0.789	326	51.988	Nucleus
FvWRKY48	41615.341	6.460	-0.982	368	62.252	Nucleus
FvWRKY49	55372.832	6.404	-0.917	505	41.078	Nucleus
FvWRKY50	64556.456	5.199	-0.713	600	46.266	-
FvWRKY51	56652.199	8.568	-0.840	519	64.220	Nucleus
FvWRKY52	49603.818	7.121	-0.616	458	54.214	Nucleus
FvWRKY53	23549.545	8.758	-0.982	213	49.765	Nucleus
FvWRKY54	28689.068	7.759	-1.021	254	52.757	-
FvWRKY55	184304.591	5.398	-0.474	1623	46.799	Nucleus
FvWRKY56	37721.405	8.497	-0.710	345	49.417	Nucleus
FvWRKY57	25326.364	9.157	-0.781	226	38.098	Nucleus
FvWRKY58	39260.246	5.225	-0.684	353	49.545	Nucleus
FvWRKY59	42415.684	6.625	-0.731	379	51.638	Nucleus
FvWRKY60	37926.638	6.435	-0.835	340	49.118	Nucleus
FvWRKY61	152199.858	6.215	-0.265	1342	42.605	Nucleus
FvWRKY62	151670.211	6.806	-0.280	1333	47.500	Nucleus
FvWRKY63	46344.290	5.122	-0.700	430	62.007	Nucleus
FvWRKY64	40957.381	6.637	-0.591	372	52.998	Nucleus

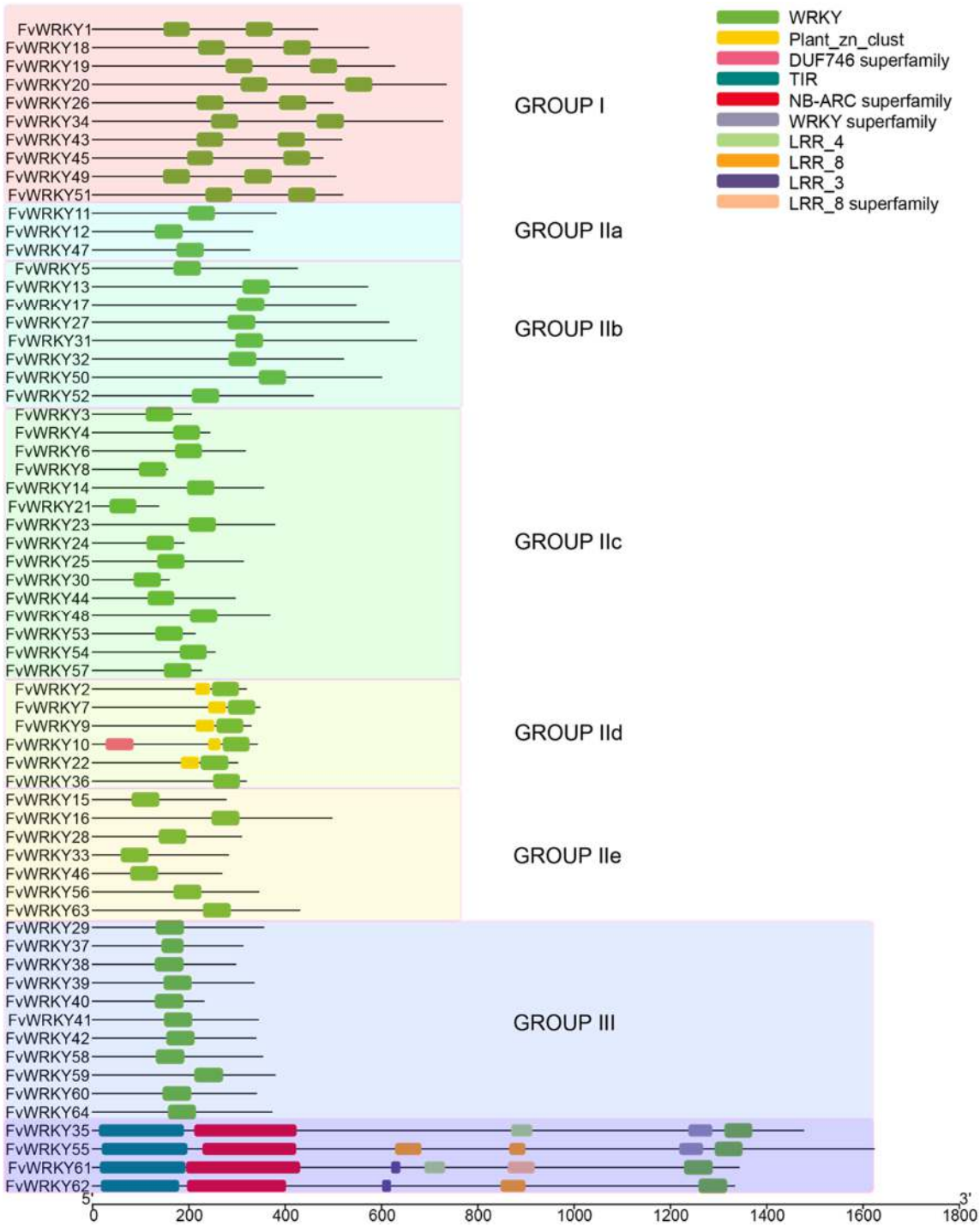


Figure 1. Protein domains (PFAM v32.0) found in FvWRKY proteins. Only largest splice forms are represented.

The recent *Fragaria x ananassa* Camarosa (Fa) Genome Assembly v1.0 and annotation v1.0.a1 was used to identify the *FaWRKY* candidates, distributed among the four diploid subgenomes which compose this allo-octoploid (Edger et al., 2019). A total of 255 FaWRKY canditates were found by HMMER, and presence of WD was confirmed in CDD (Supplementary Table S3). Homology and shared synteny with the *FvWRKY* genes were investigated using SynMap2. An additional gene (snap_masked-Fvb3-3-processed-gene-269.16) showed shared synteny with *FvWRKY20*. This Fa gene was not detected by the HMMER search due to an incomplete protein sequence in the source dataset. FGENESH (www.softberry.com) (Solovyev et al., 2006) was used to predict a revised protein sequence, using the source mRNA as input, and specific gene-finding parameters for Fv. The corrected protein sequence showed two WD and this gene was named *FaWRKY20A* (Figure 2).

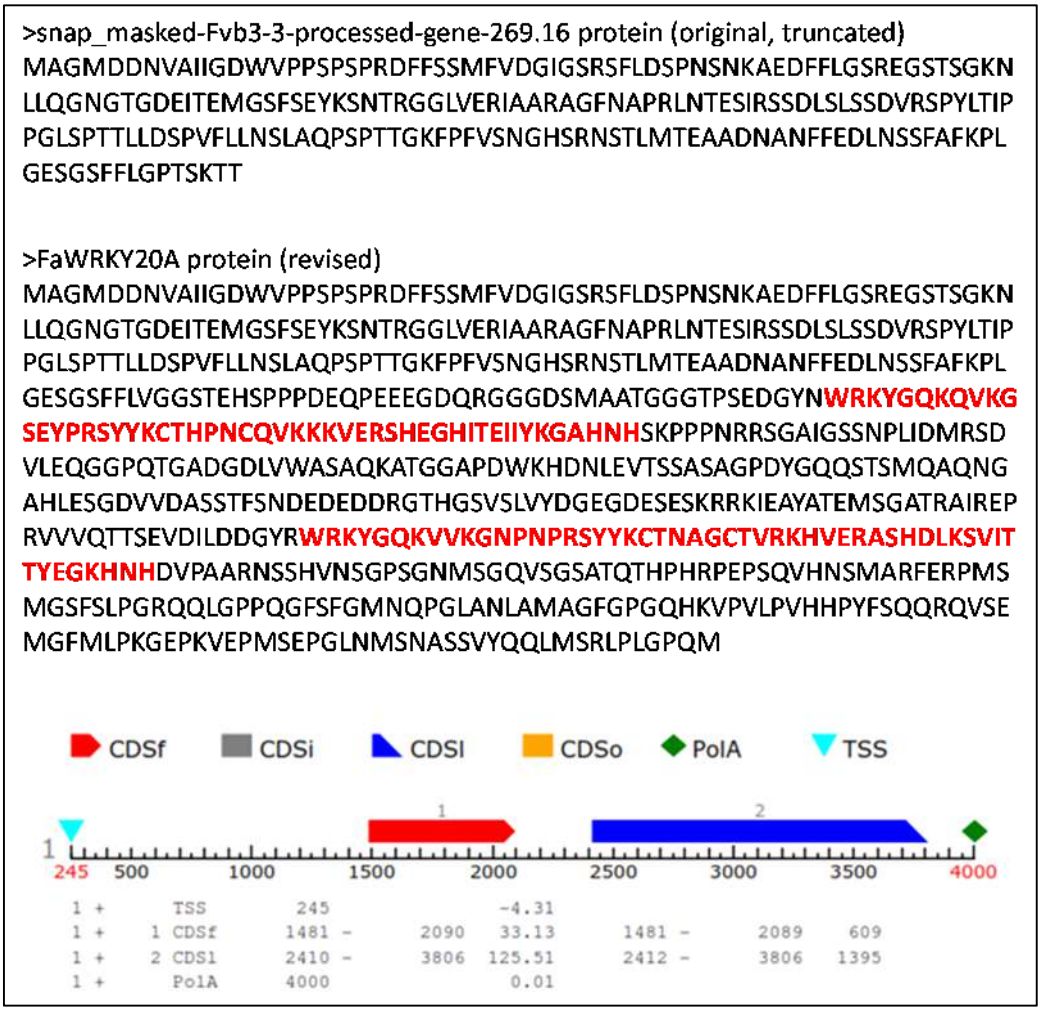


Figure 2. Original protein product of the snap_masked-Fvb3-3-processed-gene-269.16. This gene was identified as syntenic with other *FaWRKY* genes, but the originally predicted protein did not harbor WRKY domains. The source mRNA was loaded in FGENESH and a revised protein sequence was generated.

Likewise, shared synteny was found between another gene (*maker-Fvb3-1-snap-gene-3.50*) and *FvWRKY21*. A genomic DNA track containing this gene (Fvb3-1:336748..339128) was loaded in FGENESH to predict new revised mRNA and protein sequences and the new FaWRKY member was named FaWRKY21C (Figure 3). Therefore a total of 257 FaWRKY genes were found in our analysis.

>maker-Fvb3-1-snap-gene-3.50-mRNA-1 (original)
CTTCGTCCTAGTTAGCTGGCTGGCTAATTTATACATTATTTCCCAACTCAGGGTTATATATCAGAAAG
AGTGATACTTGCTTTGCTGCTCCGATCTATCTTTGCCCTCTTGAGTCTTAAGCTTACATACAAATTAC
ACAGCGAGGAAGGGAGGGTTGTATTTACAGTGTTTCTGGATAGAATCGCAGCCAAGAACAGTCG
TAGGACTAAAGTGGTGAACCATATTTTGATTCTCATCAGTCTCTCAGAAAATATACACATAAAAGGG
AAAACTGATCATCATGATGAGGATGAGGGACATGGAGAAGGAGAAGAGGAGGTGCAAGACAAT
ACCAACTCAGGCGGAAGCAACAAAGTGGTGATACCAGAAGATGGCTTCGAGTGGAGAAAATATG
GACAGAAGTTTATTAAAAATATTGGAAAGTTCAGAAGCTATTTCAAGTGCCAAAACAGCAATTGTA
GAGCCAAGAAGCGAGCCGAGTGGTCGAGCTCAGAGCCTGATAAACTTCGAGTAGTGTATGAAGG
GGTGCACAATCACACTGATAATAATGCCTCATCTGCAGGATCATCGACTTCACATTCCTCATCATCTT
TAAATCCGAGTGCTGATGCAAACCAATATGACTTGTACACACAAGTTTTTGGTAACCAATCCTCCCA
TTGGGATAAAAAAAGTTGAATCTGATAGCTGCAGCAGCAGGTTAATCCACACAGTGCATGACTACC
AGTTCTTCTGCAGCTGTTATTTACATCCCTGTTGGATAAGCCTCGCCATTTCAATATATAAGTTTAATG
CCAGTCTGTAATTGCGGTGTAAAGAAACAAGTATATCATTTGGTCCATCTATATCCGCTAATAATTTC
ATTACTGGTACAACTGATCCTGAACTCTTGGAAACAATGTCCAAGGTCAATTGACCAGCAGATATAC
AGTGTCGTATCCATATGCAAAGTTGAATAGGGTGACATACAGCATATACGTTATTGTTGAGATTGCG
CTTGCAATGCTATTGGTC
>maker-Fvb3-1-snap-gene-3.50-mRNA-1 protein (original)
MQTNMTCTHKFLVTNPPIGIKKVESDSCSSRLIHTVHDYQFFCSCYLHPCWISLAISIYKFNASL

>FaWRKY21C cDNA (revised)
ATGTCCGATAATGATTTGATTTTATATAGGAAAACTGATCATCATGATGAGGATGAGGGACATGGAG
AAGGAGAAGAGGAGGTGCAAGACAATACCAACTCAGGCGGAAGCAACAAAGTGGTGATACCAG
AAGATGGCTTCGAGTGGAGAAAATATGGACAGAAGTTTATTAAAAATATTGGAAAGTTCAGAAGC
TATTTCAAGTGCCAAAACAGCAATTGTAGAGCCAAGAAGCGAGCCGAGTGGTCGAGCTCAGAGC
CTGATAAACTTCGAGTAGTGTATGAAGGGGTGCACAATCACACTGATAATAATGCCTCATCTGCAG
GATCATCGACTTCACATTCCTCATCATCTTTAAATCCGAGTGCTGATGCAAACCAATATGACTTGTAC
ACACAAGTTTTTGGTAACCAATCCTCCCATTGGGATAAAAAAAGTTGA
>FaWRKY21C protein (revised)
MSDNDLILYRKTDHHDEDEGHGEGEEEVQDNTNSGGSNKVVIPEDGFE**WRKYGQKFIKNIGKFRSY**
FKCQNSNCRAKKRAEWSSSEPDKLRVVYEGVHNHTDNNASSAGSSTSHSSSSLNPSADANQYDLYT
QVFGNQSSHWDKKS

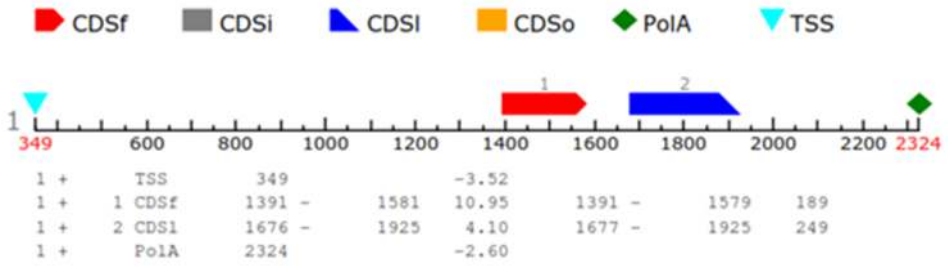


Figure 3. Original and FGENESH predicted sequences for FaWRKY21C.

Two more *FaWRKYs*, *maker-Fvb3-2-snap-gene-310.32* (FaWRKY21B) and *maker-Fvb6-2-snap-gene-308.69* (FaWRKY51A.2), were detected to have incorrect protein predicted sequences in the source dataset. Genomic DNA fragments containing each gene (Fvb3-2:31015005..31018524 and Fvb6-2:30857826..30862505, respectively) were loaded into FGENESH to recover the revised sequences of both, transcript and protein (Figures 4 and 5). All the new sequences generated were included in the successive analyses performed bellow.

>maker-Fvb3-2-snap-gene-310.32-mRNA-1 (original)
TATCAAGTAGCGAGTGGACACGTTCAGGTTTCATTGGCTGAGAATGTGATTCCGAGCCGTATAGCAG
CTGGTGCTGAGGTTATATATCAGAAAGAGCATAGAGTGATACTTGCTTTGCTGCTCCGATCTATCTCT
GCCCTCTTGAGTCTTAAGCTTACATACAAATTACACAGCGAGGAAGGGAGGGTTGTATTTACAGTGT
TTCTGGATAGAATCGCAGCCACGAGCAGTCGTAGGACTAAAGTGGTGAACCATGATGAGGGAGATG
GAGAAGGAGAAGAGGAGGTGCAAGACAACAACAATTCAGGCGGAAGCAACAAAGTGGTGATACC
AGAAGATGGCTTCGAGTGGAAAAAATATGGACAGAAGTTTATTAAAAATGTTGGAAAGTTCAGAAG
CTATTTCAAGTGCCAAAACAGCAATTGTAGAGCCAAGAAGCGAGCAGAGTGGTCGAGCTCAAAGC
CTGATAAACTTCGAGTAGTGTATGAAGGGGTGCACAATCACACTGATAATAATGCCTCATCTGCAGG
ATCATCGACATCACATTCCTCATCATCTTTAAATCCGAGTGCTGATGCAAACCAATATGACTTGTACAC
ACAAGTTTTTGGTAACCAATCCTCCCATTGGGATAAAAAAAGTTGAATCTTCTAGCTGCAGCAGCAG
GTTAATCCACACAGTGCATGACTACCAGTTCTTCTGCAGCTGTTATTTACATCCCTCTTAGATAAGCCT
CGCCATTTCAATATATAAGTTTAATGCCAGTCTGTAATTGCGGAGCAAAGAAACAAGTATATCATTTG
GTCCATCTATATCCGCTAATAATTTCATTACTGGTACAACTGATCCTGAACTCTTGGAGACAATGTCCA
AGGTCAATTAACCAGCAGATATACAAAGTGTCTTGTCCATACGCAAAGTTGAATAGGG
>maker-Fvb3-2-snap-gene-310.32 protein (original)
YQVASGHVQVSLAENVIPSRIAAGAEVIYQKEHRVILALLLRSISALLSLKLTYKLHSEEGRVVFTVFLDRIA
ATSSRRTKVVNHDEGDGEGEEEVQDNNNSGGSNKVVIPEDGFE**WKKYGQKFIKNVGKFRSYFKCQN**
SNCRAKKRAEWSSSKPDKLRVVYEGVHNHTDNNASSAGSSTSHSSSSLNPSADANQYDLYTQVFGN
QSSHWDKKS

>FaWRKY21B cDNA (revised)
ATGTCCGATTATGATTTGATTTTATATAGGAAAACTGATCATCATGATGAGGATGAGGGAGATGGAGA
AGGAGAAGAGGAGGTGCAAGACAACAACAATTCAGGCGGAAGCAACAAAGTGGTGATACCAGAA
GATGGCTTCGAGTGGAAAAAATATGGACAGAAGTTTATTAAAAATGTTGGAAAGTTCAGAAGCTAT
TTCAAGTGCCAAAACAGCAATTGTAGAGCCAAGAAGCGAGCAGAGTGGTCGAGCTCAAAGCCTGA
TAAACTTCGAGTAGTGTATGAAGGGGTGCACAATCACACTGATAATAATGCCTCATCTGCAGGATCAT
CGACATCACATTCCTCATCATCTTTAAATCCGAGTGCTGATGCAAACCAATATGACTTGTACACACAA
GTTTTTGGTAACCAATCCTCCCATTGGGATAAAAAAAGTTGA
> FaWRKY21B protein (revised)
MSDYDLILYRKTDHHDEDEGDGEGEEEVQDNNNSGGSNKVVIPEDGFE**WKKYGQKFIKNVGKFRSYF**
KCQNSNCRAKKRAEWSSSKPDKLRVVYEGVHNHTDNNASSAGSSTSHSSSSLNPSADANQYDLYTQ
VFGNQSSHWDKKS

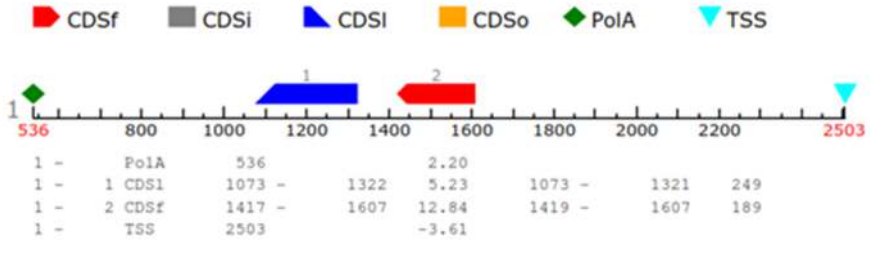


Figure 4. Original and FGENESH predicted sequences for FaWRKY21B

>maker-Fvb6-2-snap-gene-308.69-mRNA-1 (original)
TATATTGTCCATTTTTCATTTCTGCTGCGAGGACTTTCATCCCCACTTGAAATTCAGGATAGCACGAAACCCACAGAGAACA
AAACCACCCAACCGAGCAGTTGCAGAGGAAGAAAGAAAGAAGAAACAAGCATATACTTTGAAGAAGATAACATAACAG
AGTACTTCTTCTACAAAAACAGAGGATTTCAGAAACAGAGCACATCGGTAAAGCTCCCAACTTTAATGGCGGCTAAGCAC
CAAGACTCGATGTCAGCACCGCAGCCTCAGCGGCCGGTGATATCACTCCCTCCGAGGCCTTCCGCGGAGTCCCTTTTCGG
CGGCGGATCCGGAACCAGCCCAGGTCCGATGACTCTGGTCTCAAGCTTCTTCTCAGACTCAGGATTTCAGGTCCTTCTCT
CAGCTTCTTGCTGGAGCTATGGGTCACCTTTGGCTCAAGTTAGGCCTATTATGTTTGGGCAGAACCAAAACCCAGTTAGG
AATGAAGGTGCTGGGTCTGAAAATGGAGGGGAGAAGAGTTCTGGGTTTAAGCAGAGTAGGCCTATGAATTTGATGGTG
ACTCACTCGCCATTGTTTACTGTCCCACCTGGGTTGAGCCCCTCCGGGTTGCTTAACTCGCCTGCCGGGTTCTTTGCACCT
CCGAGCCCATTTGGAATATCACACCAGCAGGCCTTGGCACATGTCACTGCACTAGCTCAATCTAATTTGCAGATGCAAGCT
GAATATCAATCTTCATTACTAGCAGCTCCCATAGAGTCGCAGCCAAGTAATGCTTCCTCAATCCCCAATGAAGCTTCTCAAC
AGCAAACAACACCCTCAACATCTGAACTTGGAAGTTCCATGGTACAAACATCAGAGGCTTCTCATTCTGATAGAAAATACC
CTTCATCCCATGCCACCGATAAACCTGCAGATGATCGTTATAACTGGCGAAAATATGGGCAGAAGCAGGTTAAGGGGAGT
GAATATCCACGAAGCTACTATAAATGTACACATGTGAACTGCCCTGTCAAAAGAAAGTTGAGCGTTCACCTGCCGGCGAA
ATAACTGAGATTATCTACAAAGGGCAACATAACCATGAGCCACCTCAACCTAATAAACGTGTGAAAGATGGAGGTGGTCA
GAATGGACATATGGATTCACAGCCTAAGCTTGAAAATGGTTTACAAAGAAGAGTCGGAGATTCAACAAATCATGTGAAGC
TGTGCCTGAGAGGGATTATGAGCATACTCAAGCTGCTCCTGTGCAGTTACCAAGTGACAGTGAAGACTTGGGTGATGCG
GAAGCCAGAGAAGAAGGGGATGTTGATGAACCTAATCCAAAGAGAAGGAACATTGATGGTGTCTCATCTGAGGTAGCTT
TGCCTCACAAGACGGTGACAGAGCCTAAAATCATAGTTCAACAAGGAGTGAAGTTGATCTCCTTGATGATGGTTACAGAT
GGAGAAAGTATGGGCAGAAAGTAGTCAAAGGGAATCCTCATCCAAG
>maker-Fvb6-2-snap-gene-308.69 protein (original)
YIVHFSFLLRGLSSPLEIQDSTKPTENKTTQPSSCRGRKKEETSIYFEEDNITEYFFYKNRGFQKQSTSVKLPTLMAAKHQDSMSA
PQPQRPVISLPPRPSAESLFGGGSGTSPGPMTLVSSFFSDSGFQVLLSASCWSYGSPLAQVRPIMFGQNQNPVRNEGAGSEN
GGEKSSGFKQSRPMNLMVTHSPLFTVPPGLSPSGLLNSPAGFFAPPSPFGISHQQALAHVTALAQSNLQMQAEYQSSLLAAPI
ESQPSNASSIPNEASQQQTTPSTSELGSSMVQTSEASHSDRKYPSSHATDKPADDRYN**WRKYGQK**QVKGSEYPRSYYKCTHV
NCPVKRKLSVHLPAK

>FaWRKY51A.2 cDNA (revised)
ATGGCGGCTAAGCACCAAGACTCGATGTCAGCACCGCAGCCTCAGCGGCCGGTGATATCACTCCCTCCGAGGCCTTCCGC
GGAGTCCCTTTTCGGCGGCGGATCCGGAACCAGCCCAGCTTCTTGCTGGAGCTATGGGTCACCTTTGGCTCAAGTTAGG
CCTATTATGTTTGGGCAGAACCAAAACCCAGTTAGGAATGAAGGTGCTGGGTCTGAAAATGGAGGGGAGAAGAGTTCT
GGGTTTAAGCAGAGTAGGCCTATGAATTTGATGGTGACTCACTCGCCATTGTTTACTGTCCCACCTGGGTTGAGCCCCTCC
GGGTTGCTTAACTCGCCTGCCGGGTTCTTTGCACCTCCGAGCCCATTTGGAATATCACACCAGCAGGCCTTGGCACATGT
CACTGCACTAGCTCAATCTAATTTGCAGATGCAAGCTGAATATCAATCTTCATTACTAGCAGCTCCCATAGAGTCGCAGCCA
AGTAATGCTTCCTCAATCCCCAATGAAGCTTCTCAACAGCAAACAACACCCTCAACATCTGAACTTGGAAGTTCCATGGTA
CAAACATCAGAGGCTTCTCATTCTGATAGAAAATACCCTTCATCCCATGCCACCGATAAACCTGCAGATGATCGTTATAACT
GGCGAAAATATGGGCAGAAGCAGAAGAGTCGGAGATTCAACAAATCATGTGAAGCTGTGCCTGAGAGGGATTATGAGC
ATACTCAAGCTGCTCCTGTGCAGTTACCAAGTGACAGTGAAGACTTGGGTGATGCGGAAGCCAGAGAAGAAGGGGATG
TTGATGAACCTAATCCAAAGAGAAGGAACATTGATGGTGTCTCATCTGAGGTAGCTTTGCCTCACAAGACGGTGACAGAG
CCTAAAATCATAGTTCAACAAGGAGTGAAGTTGATCTCCTTGATGATGGTTACAGATGGAGAAAGTATGGGCAGAAAGTA
G
>FaWRKY51A.2 protein (revised)
MAAKHQDSMSAPQPQRPVISLPPRPSAESLFGGGSGTSPASCWSYGSPLAQVRPIMFGQNQNPVRNEGAGSENGGEKSSG
FKQSRPMNLMVTHSPLFTVPPGLSPSGLLNSPAGFFAPPSPFGISHQQALAHVTALAQSNLQMQAEYQSSLLAAPIESQPSNA
SSIPNEASQQQTTPSTSELGSSMVQTSEASHSDRKYPSSHATDKPADDRYN**WRKYGQK**QKSRRFNKSCEAVPERDYEHTQA
APVQLPSDSEDLGDAEAREEGDVDEPNPKRRNIDGVSSEVALPHKTVTEPKIIVQQGVKLISLMMVTDGESMGRK

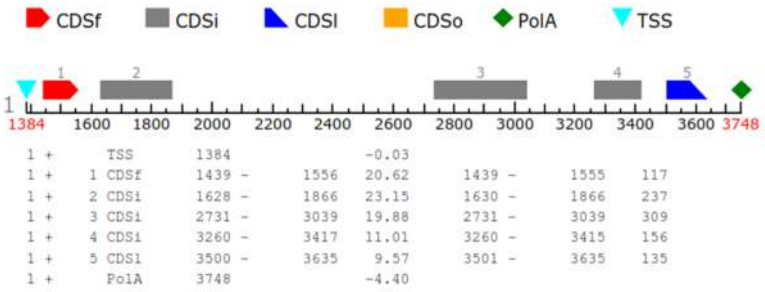


Figure 5. Original and FGENESH predicted sequences for FaWRKY51A.2

The *FaWRKY* genes were named following three criteria: foremost, due to their homology and shared synteny with the *FvWRKYs* (*FaWRKY1* to *FaWRKY64*), derived from the SynMap2 analysis; then by a letter indicating the subgenome donor (A, *F. nipponica*; B, *F. iinumae*; C, *F. viridis*; D, *F.vesca* (Table 4); and finally, numbering the gene duplications, if any (.1, .2, etc.). The final list of the 257 *FaWRKY*s is provided in Table 5. Biochemical properties and subcellular location of the FaWRKY proteins were also calculated and shown to be shared with their FvWRKY orthologs, with minor differences (data not shown). Overall, the FaWRKY family shares a high nucleotide sequence identity with their *FvWRKY* orthologs, with an average of 97.43%. Also, FaWRKY proteins in Table 5 were grouped according to their WDs, which match with their respectively FvWRKY orthologs. Furthermore, this similarity includes the additional motifs found in many of them. The number of WD modifications detected in the FaWRKY proteins is higher than the one found in the FvWRKYs, affecting FaWRKYs belonging to diverse groups, including the R protein-WRKY homologs. Many WRKY core motif variations, some of them not present in the FvWRKY proteins, as well as truncated Znf motifs, domain loss and presence of additional duplicate motifs also found in the diploid, are included. In FaWRKY51A.2, only one incomplete WD was found, lacking the Znf portion, and thus was not classified within any WRKY group. Interestingly, FaWRKY26B.2 is a chimeric protein harboring an additional Myb-like DNA-binding domain (Supplementary Table S3) as well as two unusual, modified WDs (I-N and I-C).

Also, relevant properties of the FaWRKY proteins and their predicted subcellular location are also listed in the Supplementary Table S4. As for their FvWRKY orthologs, GRAVY values indicate that FaWRKYs are hydrophilic and more likely globular-shaped, and located in the nucleus, with few exceptions.

Table 4. Letter system for subgenome assignments used in this work (adapted from Edger et al. (2019).

Diploid ancestor	Letter	Chromosome						
F. nipponica	A	Fvb1-3	Fvb2-1	Fvb3-3	Fvb4-2	Fvb5-4	Fvb6-2	Fvb7-1
F. iinumae	B	Fvb1-2	Fvb2-4	Fvb3-2	Fvb4-4	Fvb5-3	Fvb6-3	Fvb7-3
F. viridis	C	Fvb1-1	Fvb2-3	Fvb3-1	Fvb4-1	Fvb5-2	Fvb6-4	Fvb7-4
F. vesca	D	Fvb1-4	Fvb2-2	Fvb3-4	Fvb4-3	Fvb5-1	Fvb6-1	Fvb7-2

Table 5. FaWRKY family description.

Name	FaWRKY gene id	Group	WD modifications
FaWRKY1A	maker-Fvb1-3-snap-gene-6.70	I	
FaWRKY1B	snap_masked-Fvb1-2-processed-gene-48.33		
FaWRKY1C	maker-Fvb1-1-snap-gene-270.53		
FaWRKY2A	maker-Fvb1-3-snap-gene-94.39	IId	
FaWRKY2B.1	maker-Fvb1-2-augustus-gene-106.31		
FaWRKY2B.2	maker-Fvb4-4-snap-gene-200.19		incomplete Znf
FaWRKY2B.3	maker-Fvb6-3-snap-gene-191.55		CRKYGQK, incomplete Znf
FaWRKY2C.1	maker-Fvb1-1-augustus-gene-183.39		
FaWRKY2C.2	maker-Fvb1-1-augustus-gene-192.55		
FaWRKY2C.3	maker-Fvb3-1-augustus-gene-14.13		incomplete Znf
FaWRKY2D.1	maker-Fvb1-4-augustus-gene-84.44		
FaWRKY2D.2	maker-Fvb2-2-augustus-gene-143.24		WRKYRQK, incomplete Znf
FaWRKY3A	maker-Fvb1-3-snap-gene-151.44	IIc	WRKYGKK
FaWRKY3B	maker-Fvb1-2-snap-gene-163.35		WRKYGKK
FaWRKY3C	maker-Fvb1-1-snap-gene-132.18		WRKYGKK
FaWRKY3D	maker-Fvb1-4-snap-gene-131.52		WRKYGKK
FaWRKY4A	maker-Fvb1-3-augustus-gene-154.17	IIc	
FaWRKY4B	maker-Fvb1-2-augustus-gene-165.20		
FaWRKY4C	maker-Fvb1-1-augustus-gene-131.35		
FaWRKY4D	maker-Fvb1-4-augustus-gene-134.32		
FaWRKY5A	maker-Fvb1-3-augustus-gene-195.35	IIb	
FaWRKY5B	maker-Fvb1-2-snap-gene-198.47		
FaWRKY5C	maker-Fvb1-1-augustus-gene-93.38		
FaWRKY5D	maker-Fvb1-4-augustus-gene-165.28		
FaWRKY6A	maker-Fvb1-3-augustus-gene-204.36	IIc	
FaWRKY6B	maker-Fvb1-2-augustus-gene-211.23		
FaWRKY6C.1	maker-Fvb1-1-augustus-gene-16.28		
FaWRKY6C.2	maker-Fvb1-1-augustus-gene-75.42		
FaWRKY6D	maker-Fvb1-4-augustus-gene-173.37		
FaWRKY7A	maker-Fvb2-1-augustus-gene-145.32	IId	incomplete Znf
FaWRKY7B	maker-Fvb2-4-augustus-gene-166.53		
FaWRKY7C	maker-Fvb2-3-augustus-gene-103.49		
FaWRKY7D	maker-Fvb2-2-augustus-gene-65.44		
FaWRKY8A	maker-Fvb2-1-augustus-gene-207.48	IIc	WRKYGKK
FaWRKY8B	maker-Fvb2-4-augustus-gene-221.61		WRKYGKK
FaWRKY8C	maker-Fvb2-3-augustus-gene-1.22		WRKYGKK
FaWRKY8D	maker-Fvb2-2-augustus-gene-15.55		WRKYGKK

Table 5 (continued)

Name	FaWRKY gene id	Group	WD modifications
FaWRKY9A	maker-Fvb2-1-augustus-gene-222.48	IId	WQKYGQK
FaWRKY9B	maker-Fvb2-4-augustus-gene-232.43		
FaWRKY9C	maker-Fvb2-3-augustus-gene-60.42		
FaWRKY9D	maker-Fvb2-2-augustus-gene-3.56		
FaWRKY10A	maker-Fvb2-1-augustus-gene-239.50	IId	
FaWRKY10B	maker-Fvb2-4-augustus-gene-245.57		
FaWRKY10C	maker-Fvb2-3-augustus-gene-47.50		
FaWRKY11A	maker-Fvb2-1-augustus-gene-262.35	IIa	
FaWRKY11B	maker-Fvb2-4-snap-gene-265.127		
FaWRKY11C	maker-Fvb2-3-augustus-gene-23.60		
FaWRKY12A	maker-Fvb2-1-augustus-gene-262.36	IIa	
FaWRKY12C	maker-Fvb2-3-augustus-gene-23.59		
FaWRKY13B	snap_masked-Fvb3-2-processed-gene-25.34	IIb	
FaWRKY13C	maker-Fvb3-1-augustus-gene-311.52		
FaWRKY14A	maker-Fvb3-3-augustus-gene-17.65	IIc	
FaWRKY14B	augustus_masked-Fvb3-2-processed-gene-4.5		
FaWRKY14C.1	maker-Fvb3-1-augustus-gene-286.54		
FaWRKY14C.2	maker-Fvb3-1-augustus-gene-287.74		
FaWRKY14D	maker-Fvb3-4-augustus-gene-260.50		
FaWRKY15A	maker-Fvb3-3-augustus-gene-22.41	IIe	
FaWRKY15B	augustus_masked-Fvb3-2-processed-gene-40.0		
FaWRKY15C	maker-Fvb3-1-augustus-gene-280.28		
FaWRKY15D	maker-Fvb3-4-augustus-gene-255.30		
FaWRKY16A	maker-Fvb3-3-augustus-gene-42.39	IIe	
FaWRKY16B	maker-Fvb3-2-augustus-gene-58.56		
FaWRKY16C	maker-Fvb3-1-augustus-gene-261.41		
FaWRKY16D	maker-Fvb3-4-augustus-gene-237.22		
FaWRKY17A	maker-Fvb3-3-augustus-gene-126.31	IIb	
FaWRKY17B	maker-Fvb3-2-augustus-gene-145.33		
FaWRKY17C.1	maker-Fvb3-1-augustus-gene-169.33		
FaWRKY17C.2	maker-Fvb3-1-augustus-gene-173.36		
FaWRKY17D	maker-Fvb3-4-augustus-gene-165.34		
FaWRKY18A	maker-Fvb3-3-snap-gene-132.30	I	
FaWRKY18B	augustus_masked-Fvb3-2-processed-gene-154.2		
FaWRKY18C.1	maker-Fvb3-1-augustus-gene-129.27		
FaWRKY18C.2	maker-Fvb3-1-augustus-gene-147.22		
FaWRKY18D	maker-Fvb3-4-augustus-gene-158.35		

Table 5 (continued)

Name	FaWRKY gene id	Group	WD modifications
FaWRKY19A	maker-Fvb3-3-augustus-gene-259.42	I	
FaWRKY19B	maker-Fvb3-2-augustus-gene-280.44		
FaWRKY19C	maker-Fvb3-1-augustus-gene-36.35		
FaWRKY19D	maker-Fvb3-4-augustus-gene-41.39		
FaWRKY20A	snap_masked-Fvb3-3-processed-gene-269.16	I	
FaWRKY20B	maker-Fvb3-2-augustus-gene-289.30		
FaWRKY20C	maker-Fvb3-1-augustus-gene-27.44		
FaWRKY20D	maker-Fvb3-4-augustus-gene-32.34		
FaWRKY21B	maker-Fvb3-2-snap-gene-310.32	IIc	WKKYGQK
FaWRKY21C	maker-Fvb3-1-snap-gene-3.50		
FaWRKY21D	maker-Fvb3-4-augustus-gene-4.34		WKKYGQK
FaWRKY22A	maker-Fvb4-2-augustus-gene-203.21	IId	
FaWRKY22B	maker-Fvb4-4-augustus-gene-242.23		
FaWRKY22C	maker-Fvb4-1-augustus-gene-45.27		
FaWRKY22D	maker-Fvb4-3-augustus-gene-271.24		
FaWRKY23B	maker-Fvb4-4-augustus-gene-92.38	IIc	
FaWRKY23D	maker-Fvb4-3-augustus-gene-124.25		
FaWRKY24A	maker-Fvb4-2-augustus-gene-73.40	IIc	
FaWRKY24B	maker-Fvb4-4-augustus-gene-72.68		
FaWRKY24D	maker-Fvb4-3-augustus-gene-105.32		
FaWRKY25B	maker-Fvb4-4-augustus-gene-35.52	IIc	
FaWRKY25C	maker-Fvb4-1-augustus-gene-169.65		
FaWRKY25D	augustus_masked-Fvb4-3-processed-gene-36.9		
FaWRKY26A	maker-Fvb4-2-augustus-gene-31.54	I	
FaWRKY26B.1	maker-Fvb4-4-augustus-gene-34.49		
FaWRKY26B.2	maker-Fvb4-4-snap-gene-57.33		WWKYGQK (NT), WRKYRKK (CT)
FaWRKY26C	maker-Fvb4-1-augustus-gene-170.51		
FaWRKY26D	maker-Fvb4-3-augustus-gene-35.53		
FaWRKY27A	maker-Fvb4-2-snap-gene-27.61	IIb	
FaWRKY27B	maker-Fvb4-4-augustus-gene-28.60		
FaWRKY27C	maker-Fvb4-1-augustus-gene-177.49		
FaWRKY27D	maker-Fvb4-3-augustus-gene-30.43		
FaWRKY28A	maker-Fvb5-4-augustus-gene-16.45	IIe	
FaWRKY28B	maker-Fvb5-3-augustus-gene-256.41		
FaWRKY28C	maker-Fvb5-2-augustus-gene-17.32		
FaWRKY28D	maker-Fvb5-1-augustus-gene-17.52		

Table 5 (continued)

Name	FaWRKY gene id	Group	WD modifications
FaWRKY29A	maker-Fvb5-4-augustus-gene-25.53	III	
FaWRKY29B	maker-Fvb5-3-augustus-gene-251.48		
FaWRKY29C	maker-Fvb5-2-augustus-gene-27.53		
FaWRKY29D.1	maker-Fvb5-1-augustus-gene-28.49		
FaWRKY29D.2	maker-Fvb5-1-augustus-gene-32.59		
FaWRKY30A	maker-Fvb5-4-snap-gene-43.77	IIc	
FaWRKY30B	maker-Fvb5-3-snap-gene-232.60		
FaWRKY30C	maker-Fvb5-2-snap-gene-47.68		
FaWRKY30D	maker-Fvb5-1-snap-gene-49.78		
FaWRKY31A	maker-Fvb5-4-augustus-gene-67.51	IIb	
FaWRKY31B	maker-Fvb5-3-augustus-gene-193.56		
FaWRKY31C	maker-Fvb5-2-augustus-gene-85.44		
FaWRKY31D	maker-Fvb5-1-augustus-gene-95.58		
FaWRKY32A	augustus_masked-Fvb5-4-processed-gene-180.12	IIb	
FaWRKY32B	maker-Fvb5-3-augustus-gene-83.31		
FaWRKY32C	maker-Fvb5-2-augustus-gene-190.39		
FaWRKY32D	maker-Fvb5-1-augustus-gene-223.36		
FaWRKY33A	maker-Fvb5-4-snap-gene-244.45	IIe	
FaWRKY33B	maker-Fvb5-3-snap-gene-8.53		
FaWRKY33C	maker-Fvb5-2-snap-gene-244.48		
FaWRKY33D	maker-Fvb5-1-snap-gene-292.63		
FaWRKY34A	maker-Fvb6-2-augustus-gene-269.63	I	
FaWRKY34B	maker-Fvb6-3-snap-gene-427.98		
FaWRKY34C	maker-Fvb6-4-augustus-gene-26.56		
FaWRKY34D	augustus_masked-Fvb6-1-processed-gene-356.8		
FaWRKY35A	snap_masked-Fvb6-2-processed-gene-263.31	III	WTKYDQR
FaWRKY35D.1	snap_masked-Fvb6-1-processed-gene-352.19		WTKYDQR
FaWRKY35D.2	maker-Fvb6-1-augustus-gene-362.46		WTKYDQR
FaWRKY36A	maker-Fvb6-2-augustus-gene-259.49	IId	
FaWRKY36B	maker-Fvb6-3-augustus-gene-403.58		
FaWRKY36C	maker-Fvb6-4-augustus-gene-35.49		
FaWRKY36D.1	maker-Fvb6-1-augustus-gene-346.39		
FaWRKY36D.2	maker-Fvb6-1-augustus-gene-347.45		
FaWRKY37B	maker-Fvb6-3-snap-gene-362.69	III	WRK
FaWRKY37C	maker-Fvb6-4-snap-gene-67.54		WRK
FaWRKY37D	augustus_masked-Fvb6-1-processed-gene-306.17		WRK

Table 5 (continued)

Name	FaWRKY gene id	Group	WD modifications
FaWRKY38A	maker-Fvb6-2-augustus-gene-220.37	III	
FaWRKY38B	maker-Fvb6-3-augustus-gene-362.45		
FaWRKY38C	maker-Fvb6-4-augustus-gene-68.32		
FaWRKY39A	maker-Fvb6-2-augustus-gene-219.22	III	
FaWRKY40D	maker-Fvb6-1-augustus-gene-306.67	III	
FaWRKY41A	maker-Fvb6-2-augustus-gene-220.36	III	
FaWRKY41D	maker-Fvb6-1-augustus-gene-306.66		
FaWRKY42B	maker-Fvb6-3-snap-gene-362.65	III	
FaWRKY42C	maker-Fvb6-4-augustus-gene-68.34		
FaWRKY42D	maker-Fvb6-1-augustus-gene-306.65		
FaWRKY43A	augustus_masked-Fvb6-2-processed-gene-214.5	I	
FaWRKY43B.1	maker-Fvb6-3-augustus-gene-355.17		
FaWRKY43B.2	maker-Fvb6-3-augustus-gene-434.27		
FaWRKY43C	maker-Fvb6-4-augustus-gene-75.23		
FaWRKY43D	maker-Fvb6-1-augustus-gene-300.41		
FaWRKY44A	maker-Fvb6-2-augustus-gene-194.41	IIc	
FaWRKY44B	maker-Fvb6-3-augustus-gene-335.55		
FaWRKY44C	maker-Fvb6-4-augustus-gene-89.36		
FaWRKY44D	maker-Fvb6-1-augustus-gene-271.43		
FaWRKY45A	maker-Fvb6-2-augustus-gene-183.34	I	
FaWRKY45B	maker-Fvb6-3-augustus-gene-325.41		
FaWRKY45C	maker-Fvb6-4-augustus-gene-105.46		
FaWRKY45D.1	maker-Fvb6-1-augustus-gene-261.39		
FaWRKY45D.2	maker-Fvb6-1-augustus-gene-284.33		
FaWRKY46D	snap_masked-Fvb6-1-processed-gene-171.47	IIe	No heptapetide
FaWRKY47A	maker-Fvb6-2-augustus-gene-69.23	IIa	
FaWRKY47B	maker-Fvb6-3-augustus-gene-160.24		
FaWRKY47C.1	maker-Fvb6-4-augustus-gene-197.37		
FaWRKY47C.2	augustus_masked-Fvb6-4-processed-gene-215.8		
FaWRKY47D	maker-Fvb6-1-augustus-gene-164.29		
FaWRKY48A	maker-Fvb6-2-augustus-gene-359.36	IIc	
FaWRKY48B	maker-Fvb6-3-augustus-gene-82.29		
FaWRKY48C	maker-Fvb6-4-augustus-gene-286.29		
FaWRKY48D	maker-Fvb6-1-augustus-gene-89.46		
FaWRKY49A	maker-Fvb6-2-augustus-gene-344.38	I	
FaWRKY49B	maker-Fvb6-3-augustus-gene-66.19		No heptapetide (I-C)
FaWRKY49C	maker-Fvb6-4-augustus-gene-302.55		
FaWRKY49D	maker-Fvb6-1-augustus-gene-71.54		

Table 5 (continued)

Name	FaWRKY gene id	Group	WD modifications
FaWRKY50A	maker-Fvb6-2-augustus-gene-334.48	IIb	
FaWRKY50B	maker-Fvb6-3-augustus-gene-58.42		
FaWRKY50C	maker-Fvb6-4-augustus-gene-309.59		
FaWRKY50D	maker-Fvb6-1-augustus-gene-64.56		
FaWRKY51A.1	maker-Fvb6-2-augustus-gene-294.65	I	
FaWRKY51A.3	maker-Fvb6-2-snap-gene-308.69		
FaWRKY51B	maker-Fvb6-3-augustus-gene-33.67		
FaWRKY51C	maker-Fvb6-4-augustus-gene-335.56		
FaWRKY51D	maker-Fvb6-1-augustus-gene-39.52		
FaWRKY51A.2	maker-Fvb6-2-augustus-gene-307.79	NG	only I-N domain with incomplete Znf
FaWRKY52A	maker-Fvb6-2-augustus-gene-296.68	IIb	
FaWRKY52B	maker-Fvb6-3-augustus-gene-27.40		
FaWRKY52C	maker-Fvb6-4-augustus-gene-341.96		
FaWRKY52D	maker-Fvb6-1-snap-gene-33.69		
FaWRKY53A	maker-Fvb6-2-augustus-gene-278.34	IIc	
FaWRKY53B	maker-Fvb6-3-augustus-gene-9.48		
FaWRKY53C	maker-Fvb6-4-augustus-gene-16.69		
FaWRKY53D	maker-Fvb6-1-snap-gene-9.48		
FaWRKY54A.1	maker-Fvb7-1-augustus-gene-98.24	IIc	
FaWRKY54A.2	maker-Fvb7-1-augustus-gene-143.24		
FaWRKY54B	maker-Fvb7-3-snap-gene-147.40		
FaWRKY54C	maker-Fvb7-4-augustus-gene-141.43		
FaWRKY54D.1	maker-Fvb7-2-augustus-gene-91.50		
FaWRKY54D.2	maker-Fvb7-2-augustus-gene-153.32		
FaWRKY55A	maker-Fvb7-1-snap-gene-106.65	III	WREF
FaWRKY55B.1	snap_masked-Fvb7-3-processed-gene-136.20		WAKHGQK
FaWRKY55B.2	maker-Fvb7-3-augustus-gene-137.27		WAKHGQK
FaWRKY55C	maker-Fvb7-4-snap-gene-133.43		WREYDQR
FaWRKY55D	maker-Fvb7-2-augustus-gene-162.29		WREYDQR
FaWRKY56A.1	augustus_masked-Fvb7-1-processed-gene-180.2	IIe	
FaWRKY56A.2	maker-Fvb7-1-augustus-gene-192.46		
FaWRKY56B	maker-Fvb7-3-augustus-gene-111.30		
FaWRKY56C	maker-Fvb7-4-augustus-gene-106.55		
FaWRKY56D.1	maker-Fvb7-2-augustus-gene-121.31		
FaWRKY56D.2	maker-Fvb7-2-augustus-gene-187.25		

Table 5 (continued)

Name	FaWRKY gene id	Group	WD modifications
FaWRKY57A.1	maker-Fvb7-1-augustus-gene-180.37	IIc	
FaWRKY57A.2	maker-Fvb7-1-augustus-gene-192.47		
FaWRKY57B	maker-Fvb7-3-augustus-gene-111.33		
FaWRKY57C	maker-Fvb7-4-augustus-gene-106.53		
FaWRKY57D.1	maker-Fvb7-2-augustus-gene-121.33		
FaWRKY57D.2	maker-Fvb7-2-augustus-gene-187.27		
FaWRKY58A.1	maker-Fvb7-1-augustus-gene-185.48	III	
FaWRKY58A.2	maker-Fvb7-1-augustus-gene-198.36		
FaWRKY58B	maker-Fvb7-3-augustus-gene-106.33		
FaWRKY58C	maker-Fvb7-4-snap-gene-101.37		WRKYGQR
FaWRKY58D.1	maker-Fvb7-2-augustus-gene-127.28		
FaWRKY58D.2	maker-Fvb7-2-augustus-gene-192.35		
FaWRKY59D	maker-Fvb7-2-augustus-gene-251.48	III	
FaWRKY59A	maker-Fvb7-1-augustus-gene-267.38		
FaWRKY59B	maker-Fvb7-3-augustus-gene-53.56		
FaWRKY59C	maker-Fvb7-4-snap-gene-39.68		
FaWRKY60A	snap_masked-Fvb7-1-processed-gene-267.31	III	
FaWRKY60B	maker-Fvb7-3-augustus-gene-53.48		
FaWRKY60C	maker-Fvb7-4-augustus-gene-39.40		
FaWRKY60D	maker-Fvb7-2-snap-gene-251.76		
FaWRKY61B	maker-Fvb7-3-snap-gene-52.76	III	WIKYGDK
FaWRKY62A	maker-Fvb7-1-snap-gene-281.73	III	
FaWRKY62B.1	maker-Fvb7-3-snap-gene-9.78		
FaWRKY62B.2	maker-Fvb7-3-snap-gene-38.84		
FaWRKY62D	maker-Fvb7-2-snap-gene-307.51		
FaWRKY63A.1	maker-Fvb7-1-augustus-gene-291.69	IIe	
FaWRKY63A.2	maker-Fvb7-1-augustus-gene-307.58		
FaWRKY63B.1	maker-Fvb7-3-augustus-gene-0.48		
FaWRKY63B.2	maker-Fvb7-3-augustus-gene-27.73		
FaWRKY63C	maker-Fvb7-4-augustus-gene-19.70		
FaWRKY63D.1	maker-Fvb7-2-augustus-gene-279.40		
FaWRKY63D.2	maker-Fvb7-2-augustus-gene-316.52		
FaWRKY64A.1	maker-Fvb7-1-augustus-gene-293.51	III	
FaWRKY64A.2	maker-Fvb7-1-augustus-gene-309.50		
FaWRKY64B	maker-Fvb7-3-augustus-gene-26.33		
FaWRKY64C	maker-Fvb7-4-augustus-gene-14.46		
FaWRKY64D.1	maker-Fvb7-2-augustus-gene-282.49		
FaWRKY64D.2	maker-Fvb7-2-augustus-gene-319.44		

2.4.2. Genome-wide distribution and gene duplications of the strawberry WRKY family

Duplicated genes are abundant in plant genomes. These duplicates are retained after whole-genome duplication (WGD) events and have contributed to the evolution of novel functions in plants (Panchy et al., 2016). Ancient or recent WGD and polyploidy is common among the angiosperms (Lyons and Freeling, 2008;Glover et al., 2016). The homologous duplicated genes within the same genome are named paralogs, while the term homoeologs refer to the homologous genes resulting from allopolyploidy (Glover et al., 2016). We have used DAGchainer (Haas et al., 2004), implemented in the SynMap2 web-based tool (Haug-Baltzell et al., 2017), in order to identify duplicated (paralogous) *FvWRKY* genes, as well as *FaWRKY* paralogous and homeologous genes.

The *FvWRKY* genes are unevenly distributed among the seven chromosomes, with almost half of them (31 out of 64) located on chromosomes 6 and 7 (Figure 6). This is partly due to gene expansion by tandem and segmental duplications. Remarkably, 14 out of 15 group III WRKY coding genes are located on the chromosomes 6 and 7. Sixteen segmental chromosome duplications and three groups of tandem repeats containing *FvWRKY* paralogous genes were found (Table 6). All the *FvWRKY* paralogs exhibit low ω values, indicating that they are under strong negative (purifying) selection pressure as observed for other species (Wang et al., 2011). Duplicate gene pairs *FvWRKY13-50* and *FvWRKY17-50*; *FvWRKY24-53* and *FvWRKY24-30*; and *FvWRKY29-58* and *FvWRKY58-64* are sharing a common gene between, suggesting that they have evolved as a result of a two-step duplication event (Chanderbali et al., 2017). Interestingly, the two tandemly duplicated groups consisting of *FvWRKY38-39-40-41-42* and *FvWRKY59-60-61*, seems to be related to each other. The *FvWRKY61* gene, belonging to the FvRWRKY subclass, was identified as a segmental duplicate of *FvWRKY39* within a tandem repeat along with *FvWRKY59* and *-60*, pinpointing the origin of this chimeric gene as likely result of a genetic rearrangement of a group III WRKY gene and an unknown R gene leading to the formation of a novel R protein-WRKY gene. Actually, another non-WRKY gene (*FvH4_7g26100*) was identified as part of the *FvWRKY59-60-61* tandem, sharing partial sequence homology with *FvWRKY61*, as well as with a TIR-NB-LRR gene (*FvH4_7g17700*), reinforcing the idea about the origin of the R protein-like domain found in FvWRKY61 (Figure 7).

Tandem or segmental gene duplications involving the other three members of the FvRWRKY group (*FvWRKY35, -55* and *-62*) were not detected, probably because duplication-inherent mechanisms such as inversions or post-duplication events, have broken the collinear relationships among *FvWRKY* and the "donor" R genes (Lyons and Freeling, 2008). However, a candidate process already proposed for the formation and posterior expansion of the R-WRKY protein class, besides WGD events, is gene transposition (Rinerson et al., 2015). TE elements provide the capability to generate new genes by duplicating and recombining gene fragments (Sahebi et al., 2018). In a recent study about the diversification of plant immune receptors, the authors conclude that the integration of exogenous domains into Nucleotide binding leucine-rich repeat

(NLR) proteins combines both, gene duplication and interchromosomal translocation, pointing to TE elements or ectopic recombination as the most likely mechanisms (Bailey et al., 2018). Such events may also have led to the emergence of the R-WRKY protein families, independently, in several species.

Table 6. *FvWRKY* gene duplications. Nucleotide substitution values (Kn, Ks) and ω for tandemly duplicated genes were not provided by Synmap2. Genes marked with asterisks are shared by different duplicate groups.

Gene 1	Gene 2	Gene 3	Gene 4	Gene 5	Kn	Ks	ω
Segmental duplications							
FvWRKY2	FvWRKY9				0.4583	6.7082	0.0683
FvWRKY7	FvWRKY36				0.3183	1.5265	0.2085
FvWRKY13	FvWRKY50*				0.3550	2.6663	0.1331
FvWRKY14	FvWRKY48				0.4566	2.6947	0.1694
FvWRKY15	FvWRKY46				0.4185	5.7462	0.0728
FvWRKY17	FvWRKY50*				0.3908	5.7974	0.0674
FvWRKY19	FvWRKY43				0.3680	1.8100	0.2033
FvWRKY24*	FvWRKY53				0.4702	3.1511	0.1492
FvWRKY24*	FvWRKY30				0.4020	5.0096	0.0802
FvWRKY26	FvWRKY51				0.3295	3.0282	0.1088
FvWRKY27	FvWRKY31				0.7246	4.1163	0.1760
FvWRKY27	FvWRKY52				0.5880	2.9760	0.1976
FvWRKY29	FvWRKY58*				0.4700	2.2460	0.2093
FvWRKY30	FvWRKY53				0.4628	77.4216	0.0060
FvWRKY39*	FvWRKY61*				0.7849	2.7159	0.2890
FvWRKY58*	FvWRKY64				0.5863	63.4312	0.0092
Tandem repeats							
FvWRKY11	FvWRKY12				NA	NA	NA
FvWRKY38	FvWRKY39*	FvWRKY40	FvWRKY41	FvWRKY42	NA	NA	NA
FvWRKY59	FvWRKY60	FvWRKY61*			NA	NA	NA

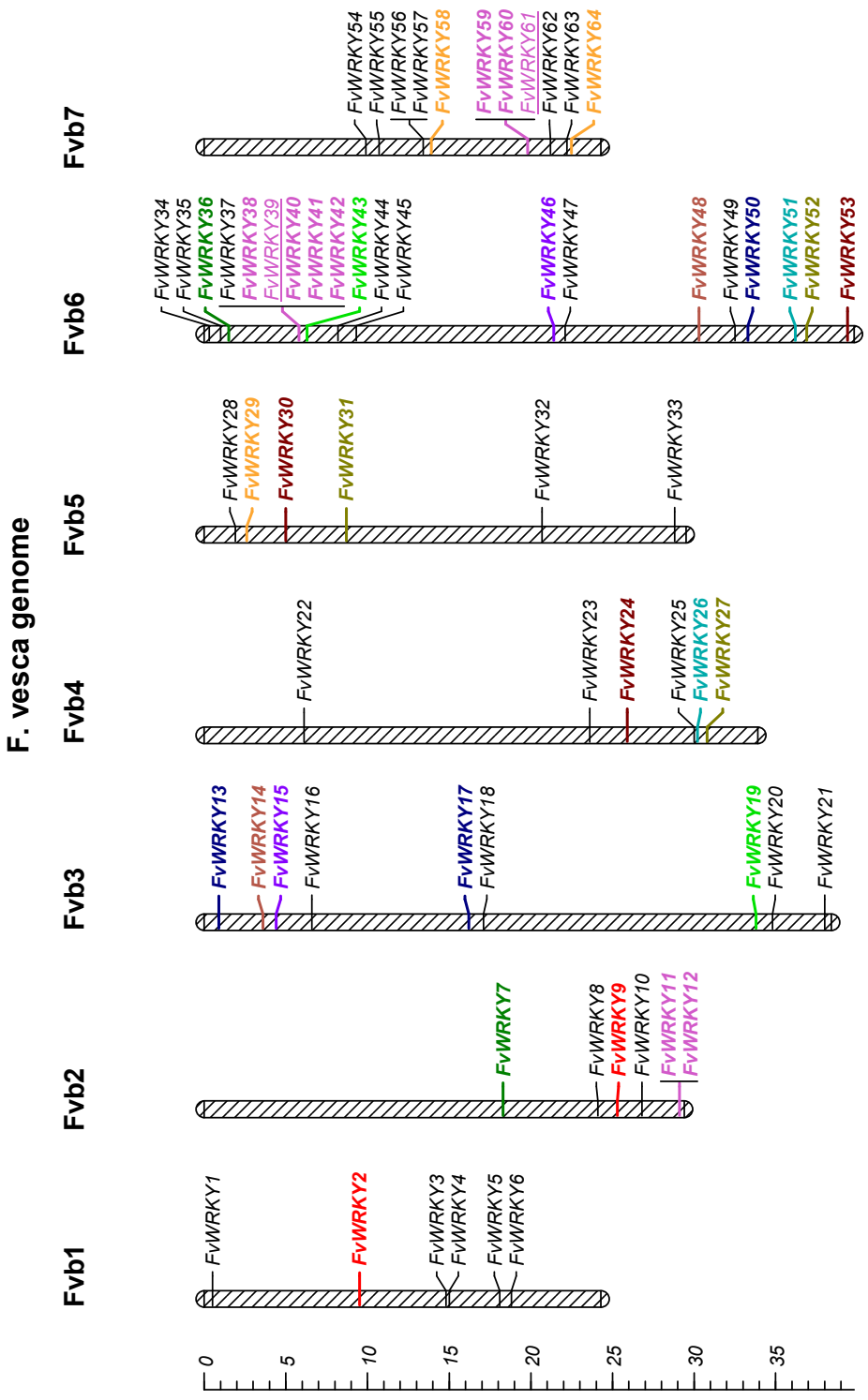


Figure 6. Chromosome mapping of the *Fragaria vesca* WRKY gene family. Segmental duplicated gene pairs (syntenic paralogs) share same colors. Tandemly duplicated gene clusters are colored in pink and underlined genes are collinear with genes aside from the cluster. Ruler size is in Megabases.

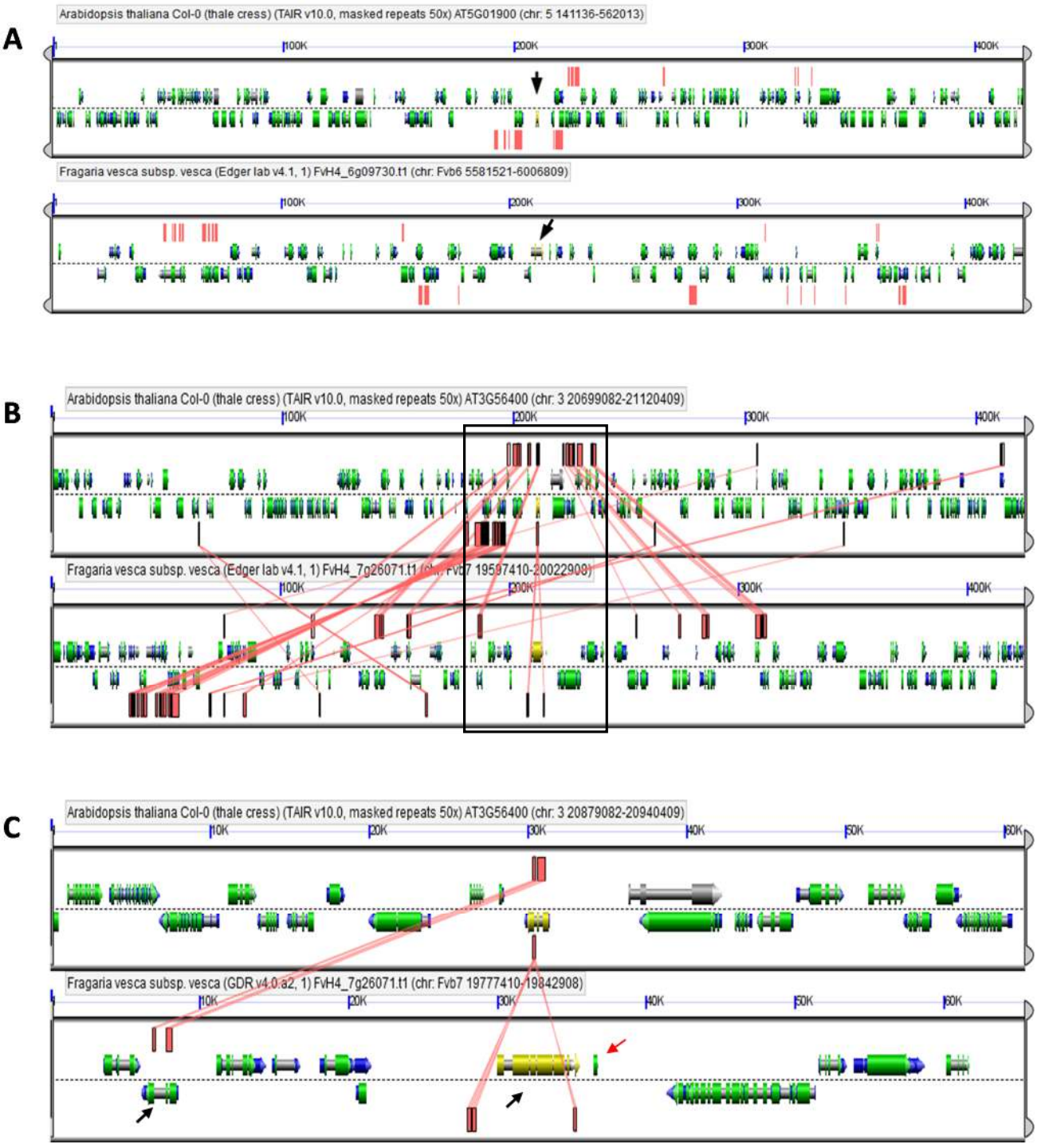


Figure 7. GEvo analyses comparing microsynteny between genomic regions from *A. thaliana* and *F. vesca*. **(A)** *FvWRKY39* and *AtWRKY62* (arrows) are located within collinear regions, despite both genes themselves don´t share hight homology. **(B)** Genomic regions containing the genes *FvWRKY60, FvWRKY61* and *AtWRKY70* (last two, in yellow). Red blocks and connectors show high scoring sequence pairs between both sequences. The region shown in C is framed. **(C)** Homology details among *FvWRKY60, FvWRKY61* (black arrows) and *AtWRKY70*. A short, non-WRKY gene (*FvH4_7g26100*, red arrow) was identified as part of the *FvWRKY59-60-61* tandem.

The octoploid Fa genome is composed by four subgenomes (A, B, C, and D) highly collinear with *F. vesca* (Figure 8) and it contains sets of homoelogous *WRKY* genes derived from the diploid ancestors (Figure 9). Genomic reorganizations and fractionation resulting from polyploidization (Tennessen et al., 2014;Edger et al., 2019), have caused some *FaWRKY* genes loss, as well as some segmental transpositions and inversions, creating collinearity breaks and hampering the identification of similar paralogous pairs as those found in the diploid. Instead, several new segmental gene duplicates were detected, as well as several non-syntenous gene duplications and triplications (Table 7). Ks values calculated for the octoploid paralogs are quite low in most cases, evidencing that gene duplications are very recent. Several *FaWRKY* paralogs show ω values greater than 1, which means that these genes may be undergoing positive selection and sub-functionalization or neo-functionalization processes might being taking place (Moore and Purugganan, 2005). Taken together, we could hypothesize that they have appeared as a result of either homoelogous exchanges, intrinsic to polyploidization (Edger et al., 2019), or post duplication events affecting single genomic features as gene transposition (Lyons and Freeling, 2008). Indeed, TEs can be activated by polyploidy and hybridization events (Vicient and Casacuberta, 2017), and play important roles in producing segmental duplications in plants and in generating changes in the genome organization and size of the hybrids, which often produce synteny breaks (Hughes et al., 2003;Bashir et al., 2018;Liu et al., 2018).

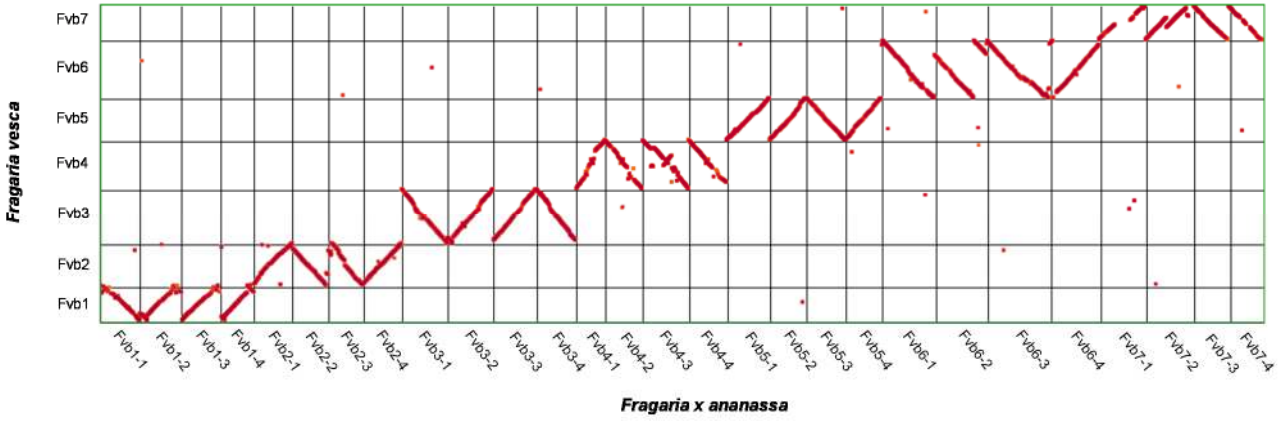


Figure 8. Macrosynteny between *F. vesca* and *F. x ananassa* genomes. The analysis reveals that genomes of both species are largely syntenic and collinear. However, several non collinear segments and genomic rearrangements are also noted.

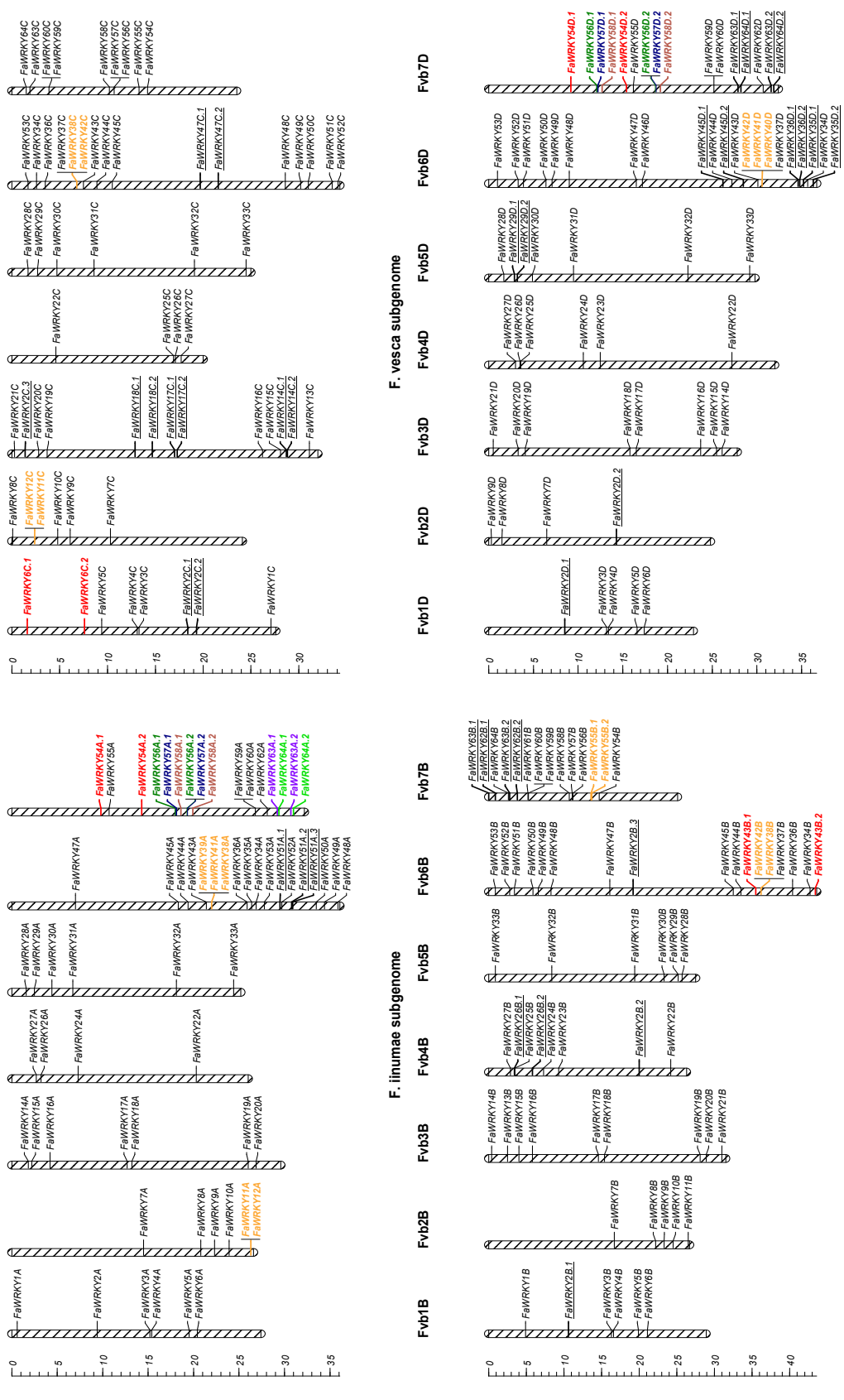


Figure 9. Chromosome mapping of the *Fragaria x ananassa* WRKY gene family, within each subgenome. Segmental duplicated gene pairs (syntenic paralogs) share same colors. Non-syntenic gene duplicates are underlined. Tandemly duplicated gene clusters are colored in orange. Ruler size is in Megabases.

Table 7. *FaWRKY* gene duplications. Nucleotide substitution values (Kn, Ks) and ω for tandemly duplicated genes were not provided by Synmap2. PAL2NAL was used to calculate Kn and Ks of the non-syntenous paralogs. ω values were not calculated if Ks<0.01.

Gene 1	Gene 2	Gene 3	Kn	Ks	ω
Segmental duplications					
FaWRKY6C.1	FaWRKY6C.2		0.0026	0.0000	NA
FaWRKY43B.1	FaWRKY43B.2		0.0017	0.0055	NA
FaWRKY54A.1	FaWRKY54A.2		0.0000	0.0039	NA
FaWRKY54D.1	FaWRKY54D.2		0.0016	0.0000	NA
FaWRKY56A.1	FaWRKY56A.2		0.0083	0.0109	0.7615
FaWRKY56D.1	FaWRKY56D.2		0.0013	0.0047	NA
FaWRKY57A.1	FaWRKY57A.2		0.0000	0.0087	NA
FaWRKY57D.1	FaWRKY57D.2		0.0022	0.0000	NA
FaWRKY58A.1	FaWRKY58A.2		0.0064	0.0111	0.5766
FaWRKY58D.1	FaWRKY58D.2		0.0012	0.0043	NA
FaWRKY63A.1	FaWRKY63A.2		0.0024	0.0000	NA
FaWRKY64A.1	FaWRKY64A.2		0.0026	0.0063	NA
Segmental, non-syntenic					
FaWRKY2D.1	FaWRKY2D.2		0.0131	0.0153	0.8562
FaWRKY2C.1	FaWRKY2C.2		0.0000	0.0000	NA
FaWRKY2C.1	FaWRKY2C.3		0.0124	0.0147	0.8482
FaWRKY2C.2	FaWRKY2C.3		0.0124	0.0147	0.8482
FaWRKY2B.1	FaWRKY2B.2		0.0134	0.0153	0.8758
FaWRKY2B.1	FaWRKY2B.3		0.0125	0.0145	0.8621
FaWRKY2B.2	FaWRKY2B.3		0.0133	0.0103	1.2913
FaWRKY14C.1	FaWRKY14C.2		0.0071	0.0279	0.2545
FaWRKY17C.1	FaWRKY17C.2		0.0021	0.0084	NA
FaWRKY18C.1	FaWRKY18C.2		0.0072	0.0069	NA
FaWRKY26B.1	FaWRKY26B.2		0.3425	0.5631	0.6082
FaWRKY29D.1	FaWRKY29D.2		0.0011	0.0000	NA
FaWRKY35D.1	FaWRKY35D.2		0.0024	0.0101	0.2376
FaWRKY36D.1	FaWRKY36D.2		0.0190	0.0322	0.5901
FaWRKY45D.1	FaWRKY45D.2		0.0000	0.0000	NA
FaWRKY47C.1	FaWRKY47C.2		0.1398	0.2113	0.6616
FaWRKY51A.1	FaWRKY51A.2		0.0400	0.0457	0.8753
FaWRKY51A.1	FaWRKY51A.3		0.0965	0.1091	0.8845
FaWRKY51A.2	FaWRKY51A.3		0.0630	0.0586	1.0751
FaWRKY62B.1	FaWRKY62B.2		0.0298	0.0269	1.1078
FaWRKY63D.1	FaWRKY63D.2		0.0022	0.0053	NA
FaWRKY63B.1	FaWRKY63B.2		0.0147	0.0132	1.1136
FaWRKY64D.1	FaWRKY64D.2		0.0025	0.0032	NA
Tandem duplications					
FaWRKY11C	FaWRKY12C		NA	NA	NA
FaWRKY11A	FaWRKY12A		NA	NA	NA
FaWRKY55B.1	FaWRKY55B.2		NA	NA	NA
FaWRKY38A	FaWRKY39A	FaWRKY41A	NA	NA	NA
FaWRKY38B	FaWRKY42B		NA	NA	NA
FaWRKY38C	FaWRKY42C		NA	NA	NA
FaWRKY40D	FaWRKY41D	FaWRKY42D	NA	NA	NA

2.4.3. Phylogenomic analysis of the strawberry WRKY family

We have used a Phylogenomic approach to study the evolution of the strawberry WRKY family and explore its potential biological implications. Shared synteny among *Fragaria vesca* (Fv), *Arabidopsis thaliana* (At) and *Vitis vinifera* (Vv) *WRKY* families was investigated using the CoGe web. SynMap3D was used to find collinear *WRKY* genes shared among the three species (Table 8). In addition, SynMap2 was used to find the collinear *WRKY* genes shared between Fv-At and Fv-Vv and not found by the SynMap3D analysis (Table 9).

Most of the FvWRKY genes are syntenous with *AtWRKY* and *VvWRKY* members of the corresponding groups. Exceptions include the R protein-WRKY members, absent in Vv, and because At and Fv members harboring similar R domains also carry different WDs as previously described by (Rinerson et al., 2015). The analysis of the ω values obtained for the syntenic At, Vv and Fv*WRKY* genes indicates that purifying selection is likely the main evolutionary driving force for this family in Rosaceae, acting against changes within the protein sequences, which modify or disrupt their functionality (Cooper and Brown, 2008). A higher number of collinear relationships are found with *VvWRKY* than *AtWRKY* genes, showing lower Ks values overall. This is probably because, unlike At, Vv has a relatively slow evolutionary rate (Jiao et al., 2012;Murat et al., 2017). Besides and as Vv, the lastest large gene expansion in the *Fragaria* lineage took place likely in the gamma (γ) whole-genome triplication, shared by the core eudicots, since there is no evidence of more recent WGD events (Jiao et al., 2012;Xiang et al., 2017;Edger et al., 2018). Two unexpected collinear relationships attract attention due to poor homology between the members involved, *FvWRKY39-AtWRKY62* (Table 8) and *FvWRKY61-AtWRKY70* (Table 9). Both strawberry genes were previously identified as part of two independent tandemly repeated gene clusters, connected each other by the FvWRKY39-FvWRKY61 segmental duplicated pair (see Table 6). The analysis of their corresponding genomic regions in GEvo showed that shared synteny between these Fv and At *WRKY* genes resulted probably from duplication-transposition of the strawberry genes into collinear regions with At (Figure 7).

Full-length WRKY proteins from the three species were aligned by MUSCLE and an unrooted phylogenetic tree was constructed and annotated using iTol to integrate both phylogenetic and collinearity relationships (Figure 10). The resulting tree distribution obtained meets the phylogenetic classification of the WRKY proteins into Groups I+IIc, IIa+IIb, IId+IIe and III (Rinerson et al., 2015).

Table 8. Collinear relationships detected by SynMap3D among *F. vesca* (Fv), *A. thaliana* (At) and *V. vinifera* (Vv) WRKY genes

Fv gene	At gene	Vv gene	Fv-At kn	Fv-Vv kn	At-Vv kn	Fv-At ks	Fv-Vv ks	At-Vv ks	Fv-At ω	Fv-Vv ω	At-Vv ω
FvWRKY1	AtWRKY32	VvWRKY35	0.4336	0.2455	0.4458	3.2006	0.9660	3.0534	0.1355	0.2541	0.1460
FvWRKY2	AtWRKY15	VvWRKY23	0.4387	0.3031	0.4687	7.8844	2.1115	5.4511	0.0556	0.1435	0.0860
FvWRKY4	AtWRKY13	VvWRKY25	0.4299	0.1330	0.4379	5.0731	1.3768	86.4744	0.0847	0.0966	0.0051
FvWRKY5	AtWRKY47	VvWRKY22	0.4755	0.4903	0.4467	8.1040	1.8793	61.5577	0.0587	0.2609	0.0073
FvWRKY11	AtWRKY18	VvWRKY10	0.4690	0.3551	0.4325	4.9340	2.0515	8.7971	0.0951	0.1731	0.0492
FvWRKY13	AtWRKY42	VvWRKY31	0.3227	0.2152	0.3648	8.5305	2.4067	13.7857	0.0378	0.0894	0.0265
FvWRKY13	AtWRKY31	VvWRKY31	0.2722	0.2152	0.2905	5.6974	2.4067	24.8433	0.0478	0.0894	0.0117
FvWRKY14	AtWRKY8	VvWRKY33	0.4453	0.2519	0.5185	3.6652	1.3246	3.5383	0.1215	0.1902	0.1465
FvWRKY14	AtWRKY28	VvWRKY33	0.4375	0.2519	0.4376	70.6736	1.3246	3.6140	0.0062	0.1902	0.1211
FvWRKY15	AtWRKY65	VvWRKY32	0.2821	0.3119	0.3197	66.8221	2.3612	2.7242	0.0042	0.1321	0.1174
FvWRKY16	AtWRKY14	VvWRKY34	0.4066	0.2316	0.3912	4.3519	1.2077	6.5210	0.0934	0.1918	0.0600
FvWRKY19	AtWRKY25	VvWRKY18	1.0319	0.2484	0.5278	73.2386	1.8339	3.4127	0.0141	0.1354	0.1547
FvWRKY24	AtWRKY75	VvWRKY3	0.2058	0.2808	0.2206	2.0126	2.0108	12.1177	0.1023	0.1396	0.0182
FvWRKY25	AtWRKY57	VvWRKY1	0.4216	0.3064	0.3232	1.7803	1.5762	3.9498	0.2368	0.1944	0.0818
FvWRKY28	AtWRKY29	VvWRKY5	0.6082	0.4662	0.5782	2.5025	2.7167	5.6765	0.2430	0.1716	0.1019
FvWRKY29	AtWRKY53	VvWRKY6	0.4414	0.3458	0.4260	27.0882	1.2382	6.0500	0.0163	0.2793	0.0704
FvWRKY36	AtWRKY39	VvWRKY43	NA	0.2878	NA	NA	1.2597	NA	NA	0.2285	NA
FvWRKY39	AtWRKY62	VvWRKY27	0.7980	0.5851	0.8162	75.1330	2.2157	6.8168	0.0106	0.2641	0.1197
FvWRKY45	AtWRKY44	VvWRKY28	NA	0.1845	NA	NA	1.0936	NA	NA	0.1687	NA
FvWRKY47	AtWRKY40	VvWRKY30	0.3695	0.1778	0.2617	4.1190	1.4743	8.9377	0.0897	0.1206	0.0293
FvWRKY49	AtWRKY1	VvWRKY39	0.6439	0.3009	0.3874	3.3571	1.3812	2.0967	0.1918	0.2179	0.1848
FvWRKY50	AtWRKY42	VvWRKY38	0.3461	0.2700	0.4888	13.2399	1.7455	68.7173	0.0261	0.1547	0.0071
FvWRKY51	AtWRKY58	VvWRKY46	0.4065	0.1952	0.5723	70.0089	2.0946	8.8434	0.0058	0.0932	0.0647
FvWRKY52	AtWRKY72	VvWRKY45	0.4021	0.2865	0.3643	8.8189	1.0447	4.5453	0.0456	0.2742	0.0801
FvWRKY52	AtWRKY72	VvWRKY54	0.4021	0.3193	0.4482	8.8189	1.3951	4.0948	0.0456	0.2289	0.1095
FvWRKY53	AtWRKY45	VvWRKY44	0.3494	0.3545	0.3156	2.4038	3.3946	80.2448	0.1454	0.1044	0.0039
FvWRKY54	AtWRKY12	VvWRKY47	0.3389	0.2036	0.2790	5.0328	1.4677	3.5763	0.0673	0.1387	0.0780
FvWRKY56	AtWRKY22	VvWRKY50	0.4646	0.4598	0.4255	1.7737	2.4831	4.9239	0.2619	0.1852	0.0864
FvWRKY57	AtWRKY43	VvWRKY49	0.6294	0.3390	NA	78.3896	1.8694	NA	0.0080	0.1813	NA
FvWRKY58	AtWRKY46	VvWRKY48	0.7190	0.3157	0.6205	3.8197	1.7930	3.0632	0.1882	0.1761	0.2026
FvWRKY58	AtWRKY46	VvWRKY6	0.7190	0.5252	0.8311	3.8197	3.0412	4.0375	0.1882	0.1727	0.2058
FvWRKY63	AtWRKY27	VvWRKY51	0.6210	0.5648	0.6093	3.8112	3.9417	19.9429	0.1629	0.1433	0.0306
FvWRKY64	AtWRKY30	VvWRKY52	0.6373	0.3070	0.6181	77.3249	1.7200	3.4856	0.0082	0.1785	0.1773

Table 9. Collinear relationships detected by SynMap2 between Fv-At and Fv-Vv WRKY genes.

Fv gene	At gene	Kn	Ks	ω	Vv gene	Kn	Ks	ω
FvWRKY2					VvWRKY55	0.217	5.1462	0.0422
FvWRKY3					VvWRKY24	0.422	2.3848	0.1770
FvWRKY6					VvWRKY21	0.3081	1.7344	0.1776
FvWRKY7					VvWRKY19	0.1086	1.1776	0.0922
FvWRKY8	AtWRKY50	0.7513	2.2441	0.3348	VvWRKY8	0.3811	2.3367	0.1631
FvWRKY9					VvWRKY55	0.2442	1.8628	0.1311
FvWRKY9					VvWRKY23	0.3968	2.5644	0.1547
FvWRKY10	AtWRKY17	0.2161	2.7874	0.0775				
FvWRKY10	AtWRKY11	0.2651	2.3031	0.1151	VvWRKY12	0.3595	2.0979	0.1714
FvWRKY13	AtWRKY6	0.3317	4.3559	0.0761	VvWRKY38	0.4378	3.0627	0.1429
FvWRKY17					VvWRKY56	0.3569	1.8493	0.1930
FvWRKY17					VvWRKY31	0.3916	3.7517	0.1044
FvWRKY18					VvWRKY57	0.1873	1.0317	0.1815
FvWRKY18					VvWRKY39	0.3973	2.3794	0.1670
FvWRKY19					VvWRKY26	0.2707	1.6606	0.1630
FvWRKY20					VvWRKY58	0.1321	0.9833	0.1343
FvWRKY21					VvWRKY17	1.5398	77.9139	0.0198
FvWRKY22					VvWRKY36	0.1829	1.6796	0.1089
FvWRKY23					VvWRKY16	0.3291	3.3779	0.0974
FvWRKY24					VvWRKY44	0.2459	2.1042	0.1169
FvWRKY26	AtWRKY4	NA	NA	NA	VvWRKY46	0.2656	2.1513	0.1235
FvWRKY27					VvWRKY2	0.3516	1.4165	0.2482
FvWRKY27					VvWRKY45	0.5206	2.2026	0.2364
FvWRKY27					VvWRKY54	0.3953	2.2364	0.1768
FvWRKY28					VvWRKY51	0.6165	4.88	0.1263
FvWRKY29					VvWRKY48	0.5488	3.6299	0.1512
FvWRKY30					VvWRKY53	0.3943	2.0936	0.1883
FvWRKY30					VvWRKY44	0.4096	7.1667	0.0572
FvWRKY30					VvWRKY3	0.4533	70.2185	0.0065
FvWRKY31					VvWRKY54	0.2908	1.3153	0.2211
FvWRKY31					VvWRKY2	0.3795	1.5322	0.2477
FvWRKY31					VvWRKY45	0.3323	2.6661	0.1246
FvWRKY33					VvWRKY20	0.2512	1.2194	0.2060
FvWRKY34					VvWRKY15	0.4327	1.7862	0.2422
FvWRKY36					VvWRKY19	0.3872	2.2051	0.1756
FvWRKY43					VvWRKY26	0.3542	1.4079	0.2516
FvWRKY44					VvWRKY29	0.2514	1.5075	0.1668
FvWRKY46	AtWRKY69	0.5424	2.211	0.2453	VvWRKY32	0.4209	2.8421	0.1481
FvWRKY48					VvWRKY33	0.2795	1.8588	0.1504
FvWRKY48					VvWRKY37	0.254	1.2441	0.2042
FvWRKY49					VvWRKY57	0.6848	3.0169	0.2270
FvWRKY50					VvWRKY56	0.3125	1.8157	0.1721
FvWRKY50					VvWRKY31	0.2802	2.2367	0.1253
FvWRKY52					VvWRKY2	0.4014	2.1746	0.1846
FvWRKY56					VvWRKY51	0.6242	63.0076	0.0099
FvWRKY61	AtWRKY70	0.6305	4.0022	0.1575				
FvWRKY63					VvWRKY5	0.7429	10.8605	0.0684
FvWRKY64					VvWRKY6	0.4127	2.5806	0.1599
FvWRKY64					VvWRKY48	0.6441	63.1558	0.0102

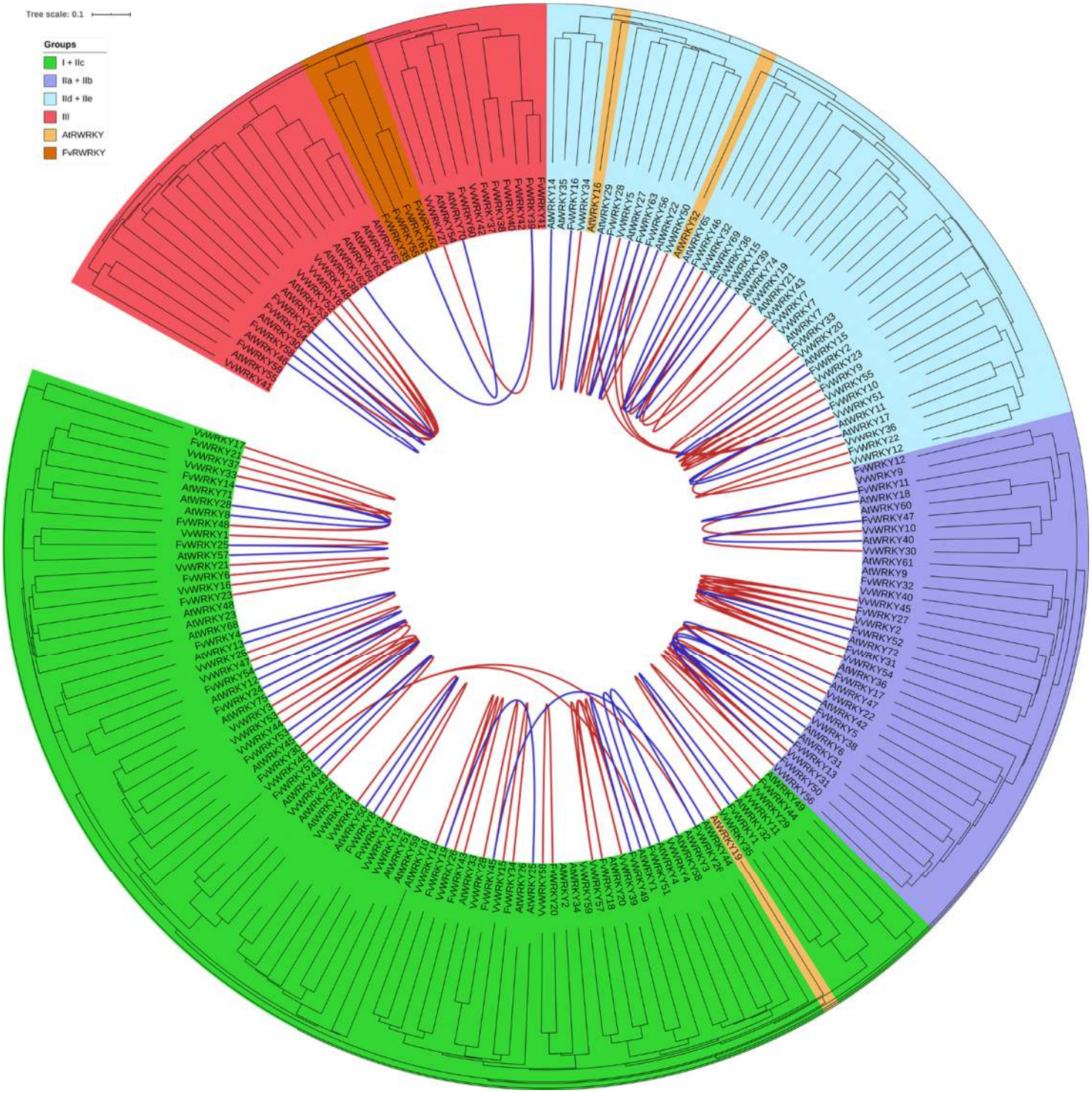


Figure 10. Phylogenetic analysis of *Fragaria vesca* (Fv), *Arabidopsis thaliana* (At) and *Vitis vinifera* (Vv) WRKY proteins. WRKY proteins are clustered into Groups I+IIc, IIa+IIb, IId+IIe and III. R protein-WRKY from Fv and At are clustered within their respective groups. The tree was inferred using the Neighbor-Joining method (1000 bootstrap replicates) and drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the p-distance method and are in the units of the number of amino acid differences per site. All positions with less than 95% site coverage were eliminated. Connecting lines represent shared synteny between Fv-Vv (red) and Fv-At (blue).

Synteny and collinearity conservation between *FvWRKY* and *FaWRKY* orthologs was also investigated with SynMap2, which found that most of the strawberry WRKYs were located within conserved blocks. Despite genomic reorganizations are widespread in the Fa genome relative to the diploid, 189 of 257 *FaWRKY* genes are syntenic and collinear with their *FvWRKY* orthologs, regardless of their subgenome origin (Table 10). The ω values show that strawberry WRKY orthologs are predominantly under purifying selection, indicating a high conservation of the WRKY gene family within the two *Fragaria* species. However, some cases of positive and possible neutral selection (due to very low Ks values and thus considered as no evolution) are also uncovered, indicating that some genes may be undergoing neo- or subfunctionalization processes. Interestingly, most of the synteny and collinearity breaks between the diploid and octoploid species are observed in those WRKY genes located in chromosomes 6 and 7 (52 out of 68), which appear to have undergone substantial genomic rearrangements in Fa, such as fractionation (loss of genomic features, like genes) and ectopic recombination with non-homologous chromosomes (Figure 11). Thus, extensive homoeologous gene losses were detected for *FaWRKY39, -40, -46,* and *-61* genes, which retain only one of the homoeologs. Several members of *FaWRKY* group III are also affected by homoeologous gene losses, perhaps because most of this group is located within these two chromosomes.

A phylogenetic tree of the diploid and octoploid strawberry WRKY proteins was constructed and annotated with the shared synteny information (Figure 12). The tree distribution obtained for the strawberry full WRKY proteins is also according with the phylogenetic classification of the WRKYs into Groups I+IIc, IIa+IIb, IId+IIe and III.

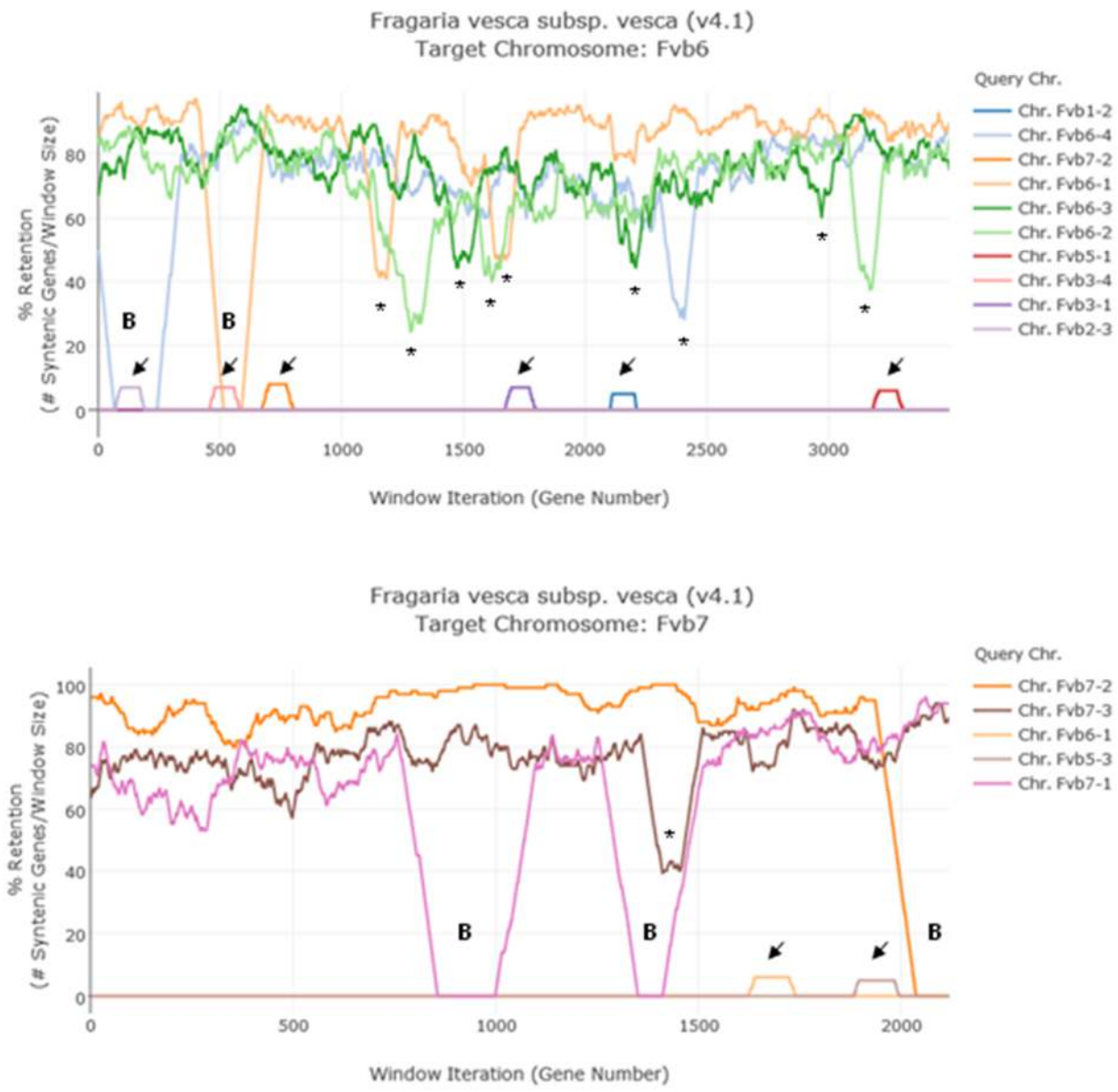


Figure 11. Gene retention and fractionation bias in Chromosomes 6 and 7 from *Fragaria vesca* and *Fragaria x ananassa* cv. Camarosa syntenic genes (1:4 syntenic depth). Whereas similar fractionation patterns are widespread in Fa chromosomes, group III FaWRKY loci, principally located in 6 and 7, seems to have been particularly affected by these events. Different regions within chromosomes have different levels of gene retention (fractionation), biased in some subgenomes (marked by "B"). Asterisks indicate areas of overfractionation. Syntenic regions with non-homologous chromosomes, potentially resulting from ectopic recombination, are indicated by arrows.

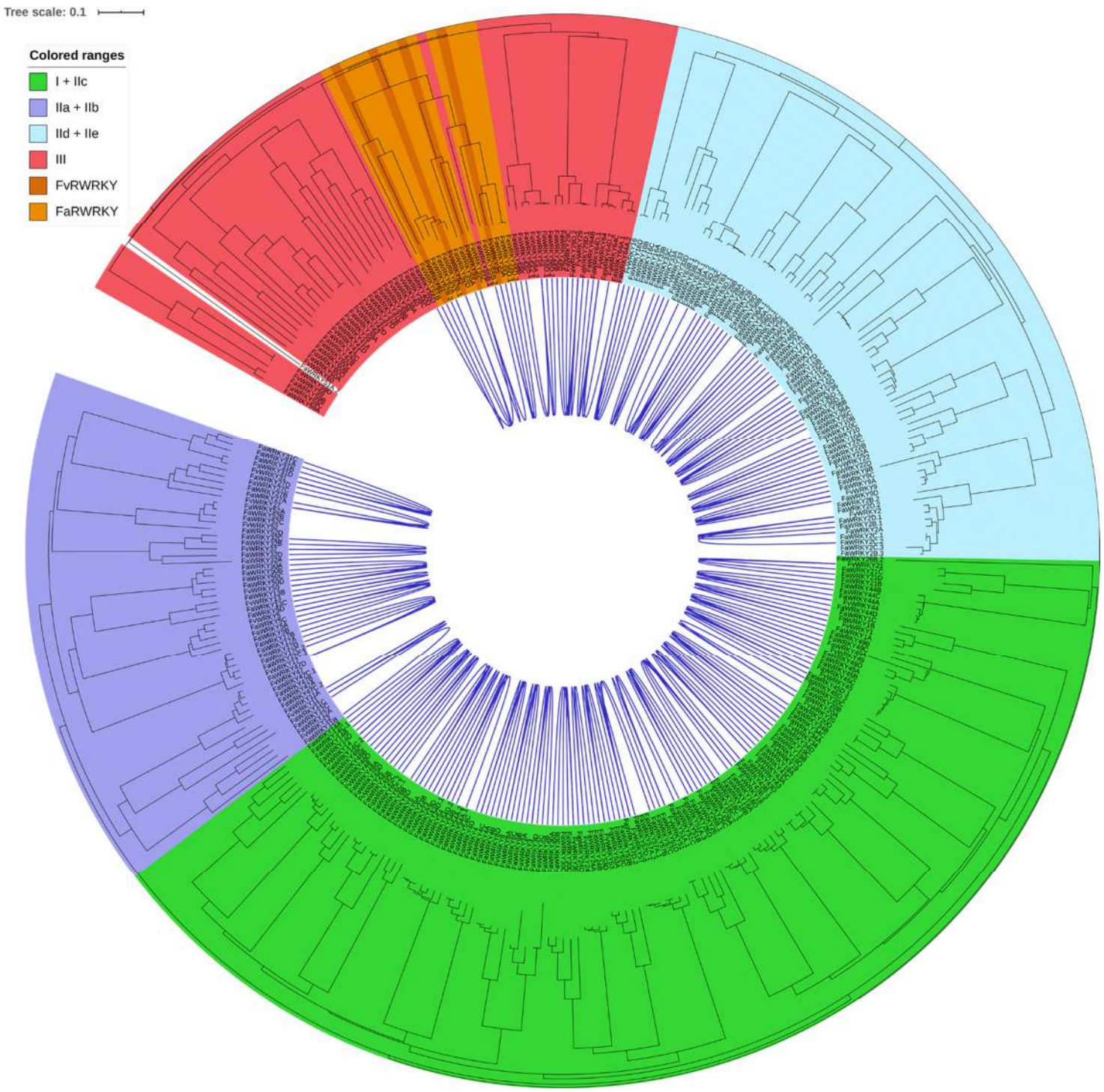


Figure 12. Phylogenetic analysis of *Fragaria vesca* (Fv) and *Fragaria x ananassa* (Fa) WRKY proteins. WRKY proteins are clustered into Groups I+IIc, IIa+IIb, IId+IIe and III. R protein-WRKY are clustered within their respective groups. The tree was inferred using the Neighbor-Joining (1000 bootstrap replicates) and drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the p-distance method and are in the units of the number of amino acid differences per site. All positions with less than 95% site coverage were eliminated. Connecting lines (blue) represent the shared synteny between Fv and Fa WRKY genes.

Table 10. Shared synteny between Fv and Fa WRKY orthologs.

FvWRKY id	FaWRKY id	Kn	Ks	ω	FvWRKY id	FaWRKY id	Kn	Ks	ω
FvWRKY1	FaWRKY1C	0.0019	0.0155	0.1226	FvWRKY15	FaWRKY15C	0.0093	0.0699	0.1330
	FaWRKY1A	0.0250	0.0458	0.5459		FaWRKY15A	0.0138	0.0857	0.1610
	FaWRKY1B	0.0165	0.0335	0.4925		FaWRKY15D	0.0032	0.0110	0.2909
FvWRKY2	FaWRKY2C.2	0.0070	0.0780	0.0897	FvWRKY16	FaWRKY16C	0.0083	0.0446	0.1861
	FaWRKY2B.1	0.0042	0.0670	0.0627		FaWRKY16B	0.0110	0.0513	0.2144
	FaWRKY2A	0.0068	0.0760	0.0895		FaWRKY16A	0.0112	0.0551	0.2033
	FaWRKY2D.1	0.0014	0.0146	0.0959		FaWRKY16D	0.0018	0.0000	NA
FvWRKY3	FaWRKY3B	0.0120	0.0430	0.2791	FvWRKY17	FaWRKY17C.2	1.2652	77.137	0.0164
	FaWRKY3D	0.0324	0.0543	0.5967		FaWRKY17A	0.0121	0.0248	0.4879
	FaWRKY3A	0.0415	0.0982	0.4226		FaWRKY17D	0.0048	0.0053	NA
	FaWRKY3C	0.0133	0.0367	0.3624		FaWRKY17B	0.0136	0.0195	0.6974
FvWRKY4	FaWRKY4C	0.0075	0.0325	0.2308	FvWRKY18	FaWRKY18C.1	0.0167	0.0391	0.4271
	FaWRKY4B	0.0090	0.0239	0.3766		FaWRKY18C.2	0.0175	0.0366	0.4781
	FaWRKY4A	0.0077	0.0300	0.2567		FaWRKY18D	0.0022	0.0000	NA
	FaWRKY4D	0.0016	0.0082	NA		FaWRKY18B	0.0170	0.0304	0.5592
FvWRKY5	FaWRKY5C	0.0190	0.0411	0.4623	FvWRKY19	FaWRKY19C	0.0120	0.0480	0.2500
	FaWRKY5B	0.0647	0.0868	0.7454		FaWRKY19A	0.0014	0.0150	0.0933
	FaWRKY5A	0.0246	0.0377	0.6525		FaWRKY19D	0.0028	0.0273	0.1026
	FaWRKY5D	0.0030	0.0000	NA		FaWRKY19B	0.0178	0.0323	0.5511
FvWRKY6	FaWRKY6C.2	0.0145	0.0254	0.5709	FvWRKY20	FaWRKY20C	0.0117	0.0487	0.2402
	FaWRKY6B	0.0114	0.0350	0.3257		FaWRKY20A	1.0272	61.567	0.0167
	FaWRKY6A	0.0148	0.0241	0.6141	FvWRKY21	FaWRKY21C	1.6135	3.8641	0.4176
	FaWRKY6D	0.0033	0.0061	NA		FaWRKY21B	0.0214	0.0548	0.3905
FvWRKY7	FaWRKY7A	0.0226	0.0938	0.2409		FaWRKY21D	0.0048	0.0179	0.2682
	FaWRKY7D	0.0015	0.0057	NA	FvWRKY22	FaWRKY22C	0.0243	0.0598	0.4064
	FaWRKY7C	0.0042	0.0256	0.1641		FaWRKY22A	0.0185	0.0822	0.2251
	FaWRKY7B	0.0043	0.0243	0.1770		FaWRKY22D	0.0060	0.0001	NA
FvWRKY8	FaWRKY8A	0.0613	0.0745	0.8228		FaWRKY22B	0.0162	0.0920	0.1761
	FaWRKY8D	NA	NA	NA	FvWRKY23	FaWRKY23D	0.0035	0.0113	0.3097
	FaWRKY8C	0.0637	0.0772	0.8251		FaWRKY23B	0.0117	0.0340	0.3441
	FaWRKY8B	0.0711	0.0704	1.0099	FvWRKY24	FaWRKY24A	0.0069	0.0414	0.1667
FvWRKY9	FaWRKY9A	0.0214	0.0567	0.3774		FaWRKY24D	0.0047	0.0000	NA
	FaWRKY9C	0.0133	0.0064	NA		FaWRKY24B	0.0087	0.0001	NA
	FaWRKY9B	0.0233	0.0346	0.6734	FvWRKY25	FaWRKY25C	0.0200	0.0716	0.2793
	FaWRKY9D	0.0963	0.1551	0.6209		FaWRKY25D	0.0072	0.0267	0.2697
FvWRKY10	FaWRKY10A	0.0026	0.0140	0.1857		FaWRKY25B	0.0233	0.0794	0.2935
	FaWRKY10C	0.0184	0.0664	0.2771	FvWRKY26	FaWRKY26A	0.0167	0.0281	0.5943
	FaWRKY10B	0.0104	0.0877	0.1186		FaWRKY26C	0.0123	0.0420	0.2929
FvWRKY11	FaWRKY11A	0.0302	0.0290	1.0414		FaWRKY26D	0.0010	0.0040	NA
	FaWRKY11C	0.0324	0.0291	1.1134		FaWRKY26B.1	0.0088	0.0417	0.2110
	FaWRKY11B	0.7131	23.2589	0.0307	FvWRKY27	FaWRKY27C	0.0286	0.0312	0.9167
FvWRKY13	FaWRKY13C	0.0156	0.0574	0.2718		FaWRKY27A	1.5030	7.9643	0.1887
	FaWRKY13B	0.0132	0.0259	0.5097		FaWRKY27D	0.0126	0.0065	NA
FvWRKY14	FaWRKY14C.1	0.0160	0.0537	0.2980	FvWRKY28	FaWRKY28D	0.0030	0.0078	NA
	FaWRKY14C.2	0.0209	0.0894	0.2338		FaWRKY28C	0.0039	0.0063	NA
	FaWRKY14B	0.0215	0.0770	0.2792		FaWRKY28B	0.0212	0.0521	0.4069
	FaWRKY14A	0.0219	0.0601	0.3644		FaWRKY28A	0.0251	0.0269	0.9331
	FaWRKY14D	0.0023	0.0000	NA					

Table 10 (continued)

FvWRKY id	FaWRKY id	Kn	Ks	ω	FvWRKY id	FaWRKY id	Kn	Ks	ω
FvWRKY29	FaWRKY29D.1	0.0000	0.0027	NA	FvWRKY49	FaWRKY49D	0.0034	0.0221	0.1538
	FaWRKY29D.2	0.0012	0.0043	NA		FaWRKY49A	0.0202	0.0405	0.4988
	FaWRKY29C	0.0167	0.0264	0.6326		FaWRKY49B	0.0250	0.0266	0.9398
	FaWRKY29B	0.0139	0.0276	0.5036		FaWRKY49C	0.0228	0.0282	0.8085
	FaWRKY29A	0.0197	0.0290	0.6793	FvWRKY50	FaWRKY50D	0.0022	0.0048	NA
FvWRKY30	FaWRKY30D	0.0032	0.0000	NA		FaWRKY50A	0.0212	0.0379	0.5594
	FaWRKY30C	0.0207	0.0247	0.8381		FaWRKY50B	0.0183	0.0294	0.6224
	FaWRKY30B	0.0142	0.0185	0.7676		FaWRKY50C	0.0175	0.0408	0.4289
	FaWRKY30A	0.0231	0.0175	1.3200	FvWRKY51	FaWRKY51D	0.0000	0.0019	NA
FvWRKY31	FaWRKY31D	0.0025	0.0000	NA		FaWRKY51B	0.0712	0.1250	0.5696
	FaWRKY31C	0.0223	0.0448	0.4978		FaWRKY51C	0.0092	0.0329	0.2796
	FaWRKY31B	0.0172	0.0379	0.4538	FvWRKY53	FaWRKY53D	0.0067	0.0055	NA
	FaWRKY31A	0.0197	0.0387	0.5090		FaWRKY53A	0.0212	0.0134	1.5821
FvWRKY32	FaWRKY32D	0.0031	0.0114	0.2719		FaWRKY53B	0.0352	0.0509	0.6916
	FaWRKY32C	0.0208	0.0382	0.5445		FaWRKY53C	0.0338	0.0235	1.4383
	FaWRKY32B	0.0124	0.0279	0.4444	FvWRKY54	FaWRKY54A.1	0.0054	0.0158	0.3418
	FaWRKY32A	0.0214	0.0612	0.3497		FaWRKY54D.1	0.0000	0.0000	NA
FvWRKY33	FaWRKY33D	0.0019	0.0032	NA		FaWRKY54D.2	0.0016	0.0000	NA
	FaWRKY33C	0.0104	0.0593	0.1754		FaWRKY54B	0.0068	0.0252	0.2698
	FaWRKY33A	0.0076	0.0658	0.1155	FvWRKY55	FaWRKY55A	0.1272	0.1055	1.2057
FvWRKY34	FaWRKY34D	0.0048	0.0163	0.2945		FaWRKY55D	0.0964	0.0955	1.0094
	FaWRKY34A	0.0049	0.0215	0.2279		FaWRKY55B.2	0.2995	0.4715	0.6352
	FaWRKY34B	0.0186	0.0403	0.4615	FvWRKY56	FaWRKY56D.1	0.0000	0.0047	NA
	FaWRKY34C	0.0159	0.0483	0.3292		FaWRKY56D.2	0.0013	0.0046	NA
FvWRKY35	FaWRKY35D.1	0.0050	0.0092	NA		FaWRKY56B	0.0122	0.0573	0.2129
	FaWRKY35D.2	0.0057	0.0100	0.5700		FaWRKY56C	0.0206	0.0502	0.4104
	FaWRKY35A	0.0194	0.0370	0.5243	FvWRKY57	FaWRKY57D.1	0.0022	0.0000	NA
FvWRKY36	FaWRKY36D.1	0.0014	0.0085	0.1647		FaWRKY57D.2	0.0000	0.0000	NA
	FaWRKY36D.2	0.0203	0.0329	0.6170		FaWRKY57B	0.0099	0.0256	0.3867
	FaWRKY36A	0.0240	0.0246	0.9756		FaWRKY57C	0.0079	0.0001	NA
	FaWRKY36B	0.0246	0.0335	0.7343	FvWRKY58	FaWRKY58D.1	0.0014	0.0028	NA
FvWRKY43	FaWRKY43A	0.0108	0.0460	0.2348		FaWRKY58D.2	0.0025	0.0000	NA
	FaWRKY43B.1	0.0130	0.0578	0.2249		FaWRKY58B	0.0225	0.0466	0.4828
	FaWRKY43B.2	0.0132	0.0505	0.2614		FaWRKY58C	0.0385	0.0947	0.4065
	FaWRKY43C	0.0069	0.0541	0.1275	FvWRKY61	FaWRKY61B	0.1658	0.4166	0.3980
FvWRKY44	FaWRKY44D	0.0000	0.0065	NA	FvWRKY62	FaWRKY62A	0.0033	0.0055	NA
	FaWRKY44A	0.0042	0.0875	0.0480		FaWRKY62B.2	0.0122	0.0306	0.3987
	FaWRKY44B	0.0108	0.0531	0.2034	FvWRKY63	FaWRKY63A.2	0.0280	0.0286	0.9790
	FaWRKY44C	0.0118	0.0691	0.1708		FaWRKY63D.1	0.0010	0.0097	NA
FvWRKY45	FaWRKY45D.1	0.0040	0.0099	NA		FaWRKY63B.2	0.0322	0.0428	0.7523
	FaWRKY45D.2	0.0040	0.0099	NA		FaWRKY63C	0.0020	0.0074	NA
	FaWRKY45A	0.0127	0.0385	0.3299	FvWRKY64	FaWRKY64A.2	0.0279	0.0279	1.0000
	FaWRKY45B	0.0086	0.0370	0.2324		FaWRKY64D.1	0.0036	0.0000	NA
	FaWRKY45C	0.0057	0.0283	0.2014		FaWRKY64B	0.0286	0.0332	0.8614
FvWRKY46	FaWRKY46D	0.1022	0.1415	0.7223		FaWRKY64C	0.0232	0.0412	0.5631
FvWRKY48	FaWRKY48D	0.0012	0.0102	0.1176					
	FaWRKY48A	0.0145	0.0527	0.2751					
	FaWRKY48B	0.0119	0.0242	0.4917					
	FaWRKY48C	0.0095	0.0192	0.4948					

2.4.4. Orthology relationships and annotation of the strawberry WRKYs

Orthologs are pairs of genes found in different species and originated by a speciation event (Glover et al., 2016). There is a widespread belief that orthologs are more prone to conserve similar functions than paralogs (Koonin, 2005). This concept, known as the "ortholog conjecture", is the basis to predict conserved gene functions. Moreover, syntenic conservation between orthologous genes is less likely to undergo positive selection, and therefore more likely to retain conserved function (Nehrt et al., 2011;Glover et al., 2016). However, inferring orthologs can be challenging, depending on the evolutionary distances and the occurrence of WGD events which produces gene duplication/loss.

Orthologs of Fv and Fa WRKY proteins were investigated using the new eggNOG 5.0 Database and hierarchically classified into orthologous groups (OGs) within several taxonomic scopes. Consistent with the high sequence conservation detected by the previous analyses, the putative FvWRKY and FaWRKY orthologs were classified into the same OGs (Table 11). Also, GO functional annotations are essentially shared by both species (Figure 13). The only exceptions were FaWRKY37D and the chimeric protein FaWRKY26B.2, which were assigned to different OGs than their respective orthologos and homoeologs. FaWRKY26B.2 was assigned to OGs of protein phosphatases 2C, with Arabidopsis orthologs ABI1 and ABI2, both related with the abscisic acid signaling pathway (Merlot et al., 2001). However, a PP2C domain is not found within this gene, which makes uncertain this function assignment. On the other hand, FaWRKY37D was classified into OGs containing non WRKY proteins of unknown function, like AT2G38570.1.

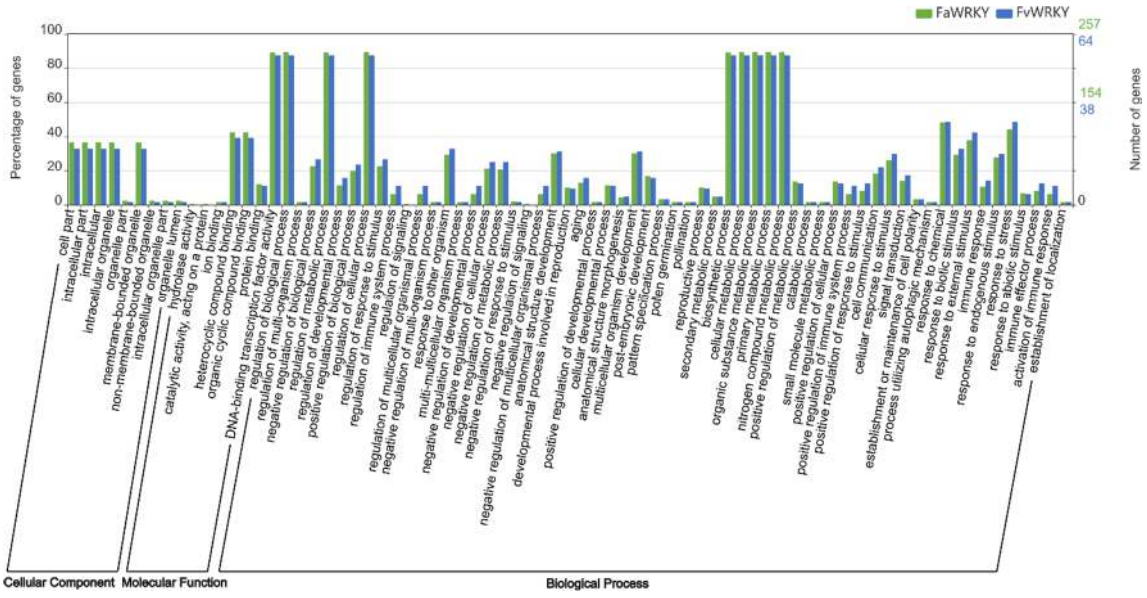


Figure 13. GO functional annotation plot comparing number and percentage of FvWRKY and FaWRKY proteins sharing same functions.

Candidate protein orthologs in several plant species can be identified in the eggNOG 5.0 Database using this hierarchical OG classification. However, we turn our attention to the homology with the AtWRKY family, which has been the best functionally characterized so far. The taxonomic scope used was the closest level to fabids containing orthologous AtWRKY proteins. Thus, up to 17 one-to-one AtWRKY orthologs were identified but one-to-many or many-to-many AtWRKY orthologs with the strawberry WRKY family were also found.

Pairwise alignment methods such as BLAST, commonly used in genome annotation pipelines, depend on the length of the compared sequences, biasing the selection of the best homolog (Emms and Kelly, 2015). Phylogenomics may be useful in most cases to distinguish between purely conserved sequences, orthologs or out-paralogs, and conserved genome regions between species, containing orthologs which might retain conserved functions. For example, several FvWRKY members (FvWRKY37, -38, -39, -40, -41, -42, -60) were found homologous to the *Arabidopsis* WRKY54/70 proteins (Table 11). The phylogenetic reconstruction using all the potential eggNOG homolog genes in several species (Supplementary Table S5) by the maximum likelihood method, shows that the strawberry *AtWRKY54/70-like* orthologous genes were originated by different rounds of gene duplication (Figure 14). Thus, the strawberry orthologs of the two Vv gamma-paralogs *VvWRKY27* and *VvWRKY42* experienced successive tamdem duplications, originating the *FvWRKY37-38-40*, *FvWRKY39-41-42* and *FvWRKY60-61* tandem gene groups. Independently, the *AtWRKY54-70* gene duplication took place likely in the more recent At-β or At-α (Tiley et al., 2016), as well as other WGD events in *Juglans*, *Malus* and *Glycine* expanded their WRKY orthologs (Schmutz et al., 2010;Velasco et al., 2010;Kim et al., 2015;Luo et al., 2015). Therefore, it is likely that several strawberry WRKY genes, homologous to *AtWRKY54-70*, share some ancestral functions of the last ones, which still today mainly show reduntant roles (Besseau et al., 2012;Chen et al., 2017a;Li et al., 2017). The two *Amborella trichopoda* orthologs included help to show that divergence between the *AtWRKY55*-like and *AtWRKY54/70*-like genes probably predates the gamma event, since it is not shared by *Amborella* (Amborella Genome, 2013). This would explain the gene *FvWRKY59,* identified as tandem duplication along with *FvWRKY60-61*, as an ancient pre-gamma paralog within this tandemly duplicated group.

Table 11. Strawberry WRKY classification into eggNOG hierarchical orthologous groups. The closest *A. thaliana* orthologs are shown.

Strawberry WRKY	eggNOG 5.0 OGs					*A. thaliana* orthologs
	root	Eukaryota	Viridiplantae	Streptophyta	Fabids	Protein
WRKY1	28JIG	2QSNR	37ND4	3GAVB	4JF9U	AtWRKY32
WRKY2	28PFA	2QU4H	37QIH	3G9MR	4JIYW	AtWRKY7, AtWRKY15
WRKY3	28JIG	2S01U	37VBB	3GIZP	4JQD6	AtWRKY50, AtWRKY59, AtWRKY51
WRKY4	28JIG	2QUMX	37JQX	3GAE0	4JMWI	AtWRKY13
WRKY5	28KH1	2QSY8	37P2H	3GC2E	4JETX	AtWRKY47
WRKY6	28JIG	2QU0Y	37KMQ	3G9WY	4JDVX	AtWRKY68, AtWRKY48, AtWRKY23
WRKY7	28PFA	2QPQX	37PI2	3GBAU	4JIED	AtWRKY21, AtWRKY39, AtWRKY74
WRKY8	28JIG	2S01U	37VBB	3GIZP	4JQ1G	AtWRKY50, AtWRKY59, AtWRKY51
WRKY9	28PFA	2QU4H	37QIH	3G9MR	4JS2R	AtWRKY7, AtWRKY15
WRKY10	28PFA	2QR3I	37NGD	3G7VS	4JNJU	AtWRKY17, AtWRKY11
WRKY11	28IW0	2QR7M	37S08	3GE86	4JFEN	AtWRKY18, AtWRKY40
WRKY12	28IW0	2RY1M	37U0Z	3GICC	4JU2B	
WRKY13	28KH1	2QSY8	37P2H	3GDXC	4JJ8V	AtWRKY42, AtWRKY31, AtWRKY6
WRKY14	28JIG	2QQYS	37SXP	3GH4E	4JHK5	AtWRKY8, AtWRKY71, AtWRKY28
WRKY15	28PFA	2QR9A	37MPS	3GH1X		AtWRKY65
WRKY16	28PFA	2QU8N	37RTH	3GFKW	4JJX9	AtWRKY35, AtWRKY14
WRKY17	28KH1	2QSY8	37P2H			AtWRKY42, AtWRKY31, AtWRKY47, AtWRKY6
WRKY18	28JIG	2QRXJ	37M8N	3GC2K	4JD7C	AtWRKY20
WRKY19	28JIG	2QRXJ	37R9D	3G8T5	4JEWW	AtWRKY25, AtWRKY33, AtWRKY26
WRKY20	28JIG	2QU86	37M7S	3GC44	4JIA1	AtWRKY2, AtWRKY34, AtWRKY10
WRKY21	28JIG	2S7GS	37X9K	3GKI0		
WRKY22	28PFA	2QR3I	37NGD	3G7VS	4JGF7	AtWRKY17, AtWRKY11
WRKY23	28JIG	2QU0Y	37KMQ	3G9WY	4JFR6	AtWRKY68, AtWRKY48, AtWRKY23
WRKY24	28JIG	2RZAJ	37UFY	3GI6R	4JTU2	AtWRKY75
WRKY25	28JIG	2QSI6	37QAE	3GB61	4JJEW	AtWRKY57
WRKY26	28JIG	2QTWW	37K4Y	3GA0G	4JH8P	AtWRKY4, AtWRKY58, AtWRKY3
WRKY27	28KH1	2QVE0	37PNG	3G9A1	4JT1F	AtWRKY61, AtWRKY72
WRKY28	28PFA	2RZFX	37UXW	3GIPZ	4JPQX	AtWRKY29
WRKY29	2CMA2	2QPRM	37KWZ	3GGNR	4JVRN	AtWRKY30, AtWRKY53, AtWRKY41
WRKY30	28JIG	2RZAJ				AtWRKY45, AtWRKY75
WRKY31	28KH1	2QVE0	37PNG	3G9A1	4JIV6	AtWRKY61, AtWRKY72
WRKY32	28KH1	2QU92	37KWR	3G8VF	4JH50	AtWRKY9

Table 11 (continued)

Strawberry WRKY	eggNOG 5.0 OGs					*A. thaliana* orthologs
	root	Eukaryota	Viridiplantae	Streptophyta	Fabids	Protein
WRKY33	28PFA	2R27Y	37TNP	3GFNI	4JNXU	
WRKY34	28JIG	2QRFS	37TAB	3G8SD	4JE2R	
WRKY35	28XIB	2R4BG	37T5Z	3GHA6	4JRR9	
WRKY36	28PFA	2QPQX	37PI2	3GBAU	4JIED	AtWRKY21, AtWRKY74, AtWRKY39
WRKY37	2A6Y9	2RYCZ	37UB3	3GGNB	4JPDM	AtWRKY54, AtWRKY70
WRKY38	2A6Y9	2RYCZ	37UB3	3GGNB	4JPDM	AtWRKY54, AtWRKY70
WRKY39	2A6Y9	2RYCZ	37UB3	3GGNB	4JPDM	AtWRKY54, AtWRKY70
WRKY40	2A6Y9	2RYCZ	37UB3	3GGNB	4JPDM	AtWRKY54, AtWRKY70
WRKY41	2A6Y9	2RYCZ	37UB3	3GGNB	4JPDM	AtWRKY54, AtWRKY70
WRKY42	2A6Y9	2RYCZ	37UB3	3GGNB	4JPDM	AtWRKY54, AtWRKY70
WRKY43	28JIG	2RUW1	37TD4	3GHHB	4JS18	
WRKY44	28PFA	2R1XV	37T31	3GFKQ	4JE8K	AtWRKY49
WRKY45	28JIG	2QS6E	37PQH	3GFV8	4JEN2	AtWRKY44
WRKY46	28PFA	2QR9A	37MPS	3GH1X	4JTIJ	AtWRKY65
WRKY47	28IW0	2QR7M	37S08	3GE86	4JFEN	AtWRKY18, AtWRKY40
WRKY48	28JIG	2QQYS	37SXP	3GH4E	4JM4U	AtWRKY8, AtWRKY71, AtWRKY28
WRKY49	28JIG	2QV11	37KFU	3GA1G	4JKUS	AtWRKY1
WRKY50	28KH1	2QSY8	37P2H	3GDXC	4JMK4	AtWRKY42, AtWRKY31, AtWRKY6
WRKY51	28JIG	2QTWW	37K4Y	3GA0G	4JHKC	AtWRKY4, AtWRKY58, AtWRKY3
WRKY52	28KH1	2QVE0	37PNG	3G9A1	4JT7E	AtWRKY61, AtWRKY72
WRKY53	28JIG	2RZAJ	37UFY			AtWRKY45, AtWRKY75
WRKY54	28PFA	2QTRJ	37M40	3GI5B	4JSJ3	AtWRKY12
WRKY55	28XIB	2R4BG	37T5Z	3GHA6	4JRR9	
WRKY56	28PFA	2QW38	37P6B	3GGIY	4JHBF	AtWRKY22
WRKY57	28JIG	2RXZS	37U8Z	3GIAP	4JPVH	AtWRKY56, AtWRKY24, AtWRKY43
WRKY58	2CMA2	2QPRM	37KWZ	3GGNR	4JSA1	AtWRKY30, AtWRKY53, AtWRKY41
WRKY59	2A6Y9	2RSM6	37MH7	3GBKR	4JNXI	AtWRKY55
WRKY60	2A6Y9	2RYCZ	37TRF	3GIBM	4JP06	AtWRKY54, AtWRKY70
WRKY61	28XIB	2R4BG	37T5Z	3GHA6	4JRR9	
WRKY62	28XIB	2R4BG	37T5Z	3GHA6	4JRR9	
WRKY63	28PFA	2QW38	37P6B	3GEZD	4JP2P	AtWRKY27
WRKY64	2CMA2	2QPRM	37KWZ	3G9RK	4JF9W	AtWRKY30, AtWRKY53, AtWRKY41
FaWRKY26B.2	KOG4658	KOG4658	37M3H	3GA8T	4JEBM	ABI1, ABI2
FaWRKY37D	2CN2Y	2QTM7	37HU7	3G9PU	4JN92	AT2G38570

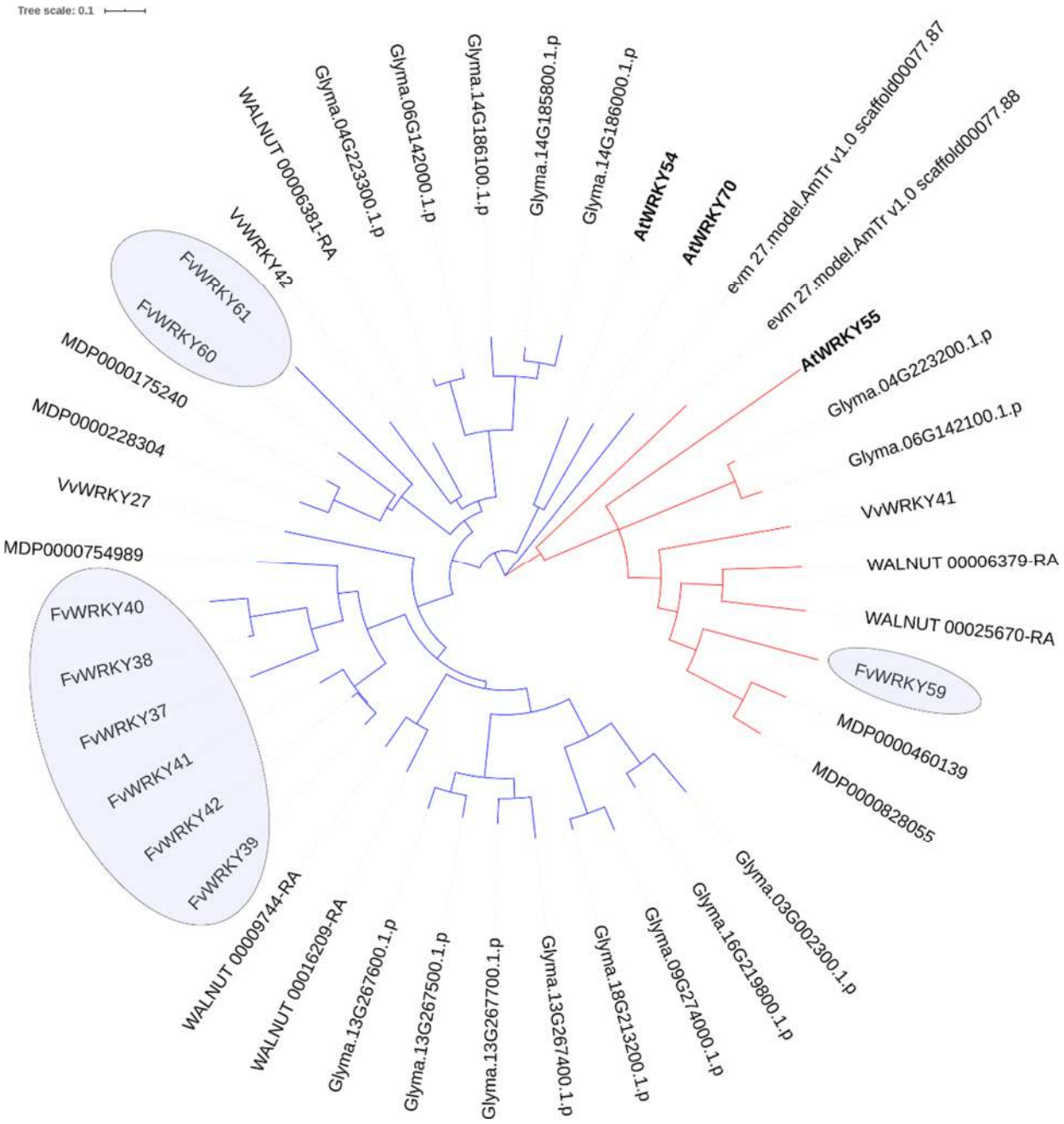


Figure 14. Phylogenetic tree of the *AtWRKY55* and *AtWRKY54/70* orthologous genes (red and blue branches, respectively) in strawberry (Fv), soybean (Glyma), walnut, apple (MD), grapevine (Vv) and *Amborella (AmTr)*. Strawberry orthologs to AtWRKY54/70 paralogs underwent tandem duplications after the gamma hexaploidization, and independently to the more recent WGD events in other species. The evolutionary history was inferred using the Maximum Likelihood method (100 bootstrap) with optimized parameters (TN+G+I). The tree is drawn to scale, with lengths of branches repressenting the number of substitutions per site. Analyses were conducted in MEGA7.

2.4.5. Expression analysis of *FvWRKY* genes in strawberry tissues and different development stages and growth conditions.

Most of strawberry WRKYs have not been yet functionally characterized nor fully described at the expression level, with few exceptions (Encinas-Villarejo et al., 2009;Wei et al., 2018;Higuera et al., 2019). Consequently, to gain insight about their possible biological roles, the expression pattern of the complete *FvWRKY* family was analysed in several tissues and developmental stages (Hollender et al., 2012) considering the transcript expression values collected from previous research (Li et al., 2019) (Supplemental Table S6).

The analysis revealed that four genes, *FvWRKY10, -2, -7* and *-33* exhibit high transcript accumulation in all samples, except pollen, suggesting that they should play important roles in most phases of strawberry physiology. Both, transcript abundance and expression patterns of the remaining *FvWRKY* genes varies largely in different vegetative and reproductive tissues, as well as during receptacle ripening and pathogen challenge (Figure 15). For example, pollen shows very low transcript accumulation of most FvWRKYs, except for *FvWRKY20* and *-36*, stronghly supporting that they could play essential roles in pollen function, perhaps being stored in ribonucleoprotein particles, as suggested by Hafidh et al. (2018) (Hafidh et al., 2018). In fact, two putative FvWRKY20 orthologs in *Arabidopsis*, AtWRKY34 and AtWRKY2, are involved in pollen development, germination, and pollen tube growth (Guan et al., 2014).

On the contrary, several *FvWRKY* genes changed their transcripts abundance or expression profiles during ripening in receptacle from two botanical forms of *F. vesca*, Ruegen (F7–4) and Yellow Wonder 5AF7, which produce fruits with red or yellow flesh and skin, respectively. However, the changes detected in the expression pattern from 15 days to 22 days post-anthesis for *FvWRKY2, -23, -29, -36, -47* and *-51,* were similar in both accessions, pointing out that these genes may play key roles in the strawberry fruit ripening process.

Interestingly, a greater number of *FvWRKY* genes (49 out of 64) changed their expression pattern in root infected by the hemibiotrophic oomycete *Phytophthora cactorum*, with transcript abundance differences that reflect both positive and negative regulation in response to this pathogen (Toljamo et al., 2016). Notorious upregulation was detected for orthologous genes of well-known AtWRKY TFs involved in plant defense responses (see Table 11). Among them, FvWRKY3 and FvWRKY8 share homology to AtWRKY50/51, which are known to mediate SA- and JA signaling enhancing resistance to some pathogens but increasing susceptibility to others (Gao et al., 2011;Hussain et al., 2018). Remarkable up regulation was found for *FvWRKY24*, *FvWRKY30* and *FvWRKY53*. These three genes are orthologs of AtWRKY75 and AtWRKY45. AtWRKY75, and its also *F. ananassa* ortholog FaWRKY24 (previously reported as FaWRKY1) have been described as positive and negative regulators of plant defense in *Arabidopsis* and *F. ananassa*, respectively (Encinas-Villarejo et al., 2009;Higuera et al., 2019;Chen et al., 2021). Also, FvWRKY29 and FvWRKY58 share

orthologous group with AtWRKY53, which positively modulate SAR in *Arabidopsis* (Eulgem, 2006;Wang et al., 2006) but it has been imply in other physiological processes like senescence and drough tolerance (Sun and Yu, 2015;Phukan et al., 2016). Similarly, FvWRKY38, FvWRKY42 and FvWRKY60 share homology with AtWRKY70 and AtWRKY54, its closest homolog, which seem to regulate the balance between SA- and JA-dependent defense pathways, acting as negative regulator of SA biosynthesis (Li et al., 2017). Also, it is worth mentioning FvWRKY19, FvWRKY43 and FvWRKY47 that share orthologous group with AtWRKY25, AtWRKY33 and AtWRKY18/40, respectively. These last ones, are WRKY major factors predicted to function as important hubs within the WRKY network of plant defense in *Arabidopsis* (Zheng et al., 2007;Birkenbihl et al., 2018). All in all, these data suggest an outstanding involvement of the WRKY TF family in the woodland strawberry defense response.

A closer look at the transcription pattern (Figure 16) reveals substantial differences in the transcription level of most paralogous pairs, as well as in their expression profiles, which suggest that they are not fully redundant and may have undergone functional divergence, neo- or sub-fuctionalization (Roulin et al., 2013). Thus, duplicate genes *FvWRKY2-9, FvWRKY7-36, FvWRKY13-50, FvWRKY17-50, FvWRKY19-43, FvWRKY24-30, FvWRKY24-53, FvWRKY29-58, FvWRKY58-64* in receptacle ripening; *FvWRKY2-9, FvWRKY11-12, FvWRKY15-46* in root infection; and members of the tandem duplications *FvWRKY38-39-40-41-42* and *FvWRKY59-60-61* in both, receptacle ripening and root infection, show very different expression patterns.

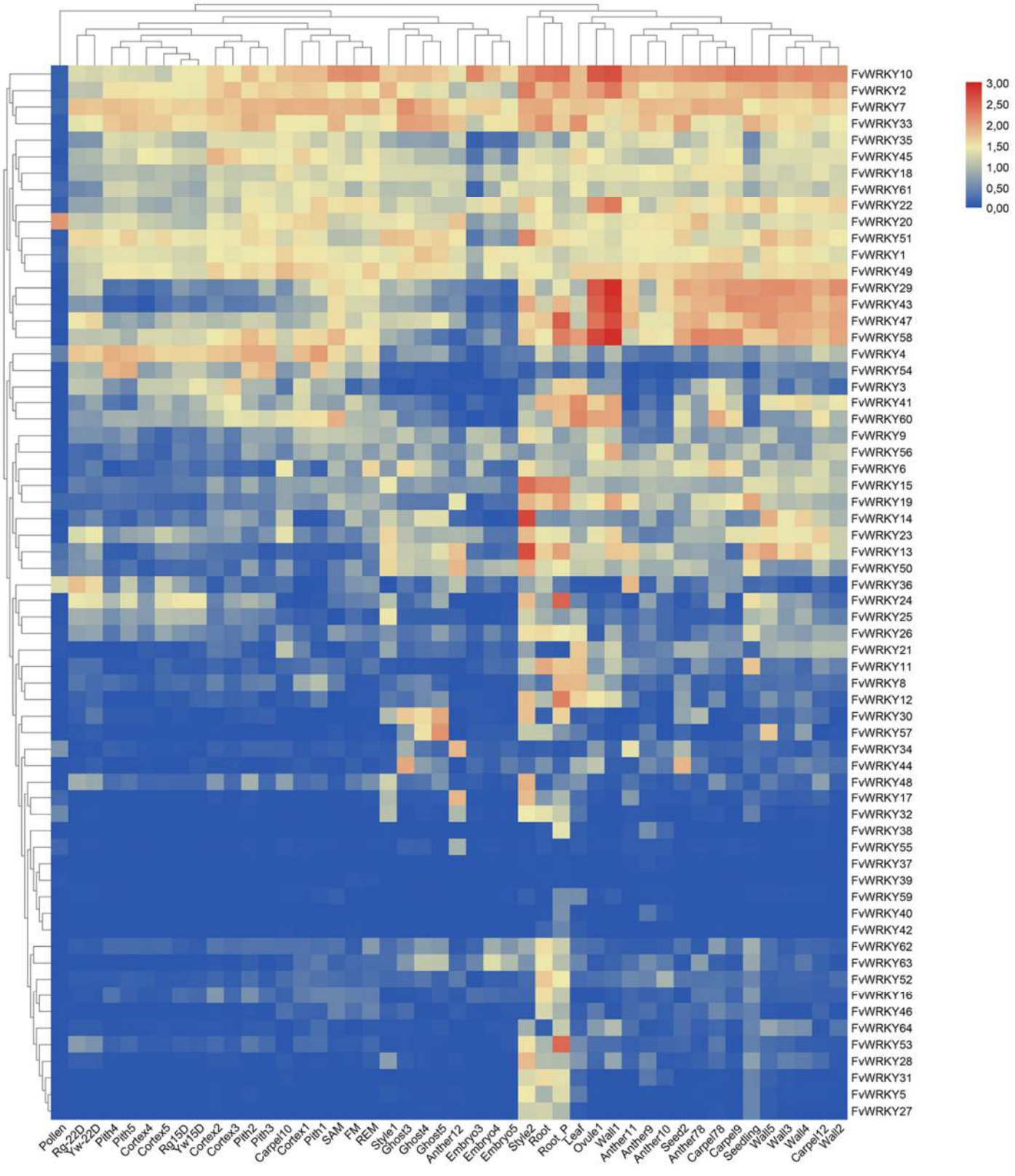

FvWRKY10
FvWRKY2
FvWRKY7
FvWRKY33
FvWRKY35
FvWRKY45
FvWRKY18
FvWRKY61
FvWRKY22
FvWRKY20
FvWRKY51
FvWRKY1
FvWRKY49
FvWRKY29
FvWRKY43
FvWRKY47
FvWRKY58
FvWRKY4
FvWRKY54
FvWRKY3
FvWRKY41
FvWRKY60
FvWRKY9
FvWRKY56
FvWRKY6
FvWRKY15
FvWRKY19
FvWRKY14
FvWRKY23
FvWRKY13
FvWRKY50
FvWRKY36
FvWRKY24
FvWRKY25
FvWRKY26
FvWRKY21
FvWRKY11
FvWRKY8
FvWRKY12
FvWRKY30
FvWRKY57
FvWRKY34
FvWRKY44
FvWRKY48
FvWRKY17
FvWRKY32
FvWRKY38
FvWRKY55
FvWRKY37
FvWRKY39
FvWRKY59
FvWRKY40
FvWRKY42
FvWRKY62
FvWRKY63
FvWRKY52
FvWRKY16
FvWRKY46
FvWRKY64
FvWRKY53
FvWRKY28
FvWRKY31
FvWRKY5
FvWRKY27
3,00
2,50
2,00
1,50
1,00
0,50
0,00
Pollen
Rg-22D
Yw-22D
Pith4
Pith5
Cortex4
Cortex5
Rg15D
Yw15D
Cortex2
Cortex3
Pith2
Pith3
Carpel10
Cortex1
Pith1
SAM
FM
REM
Style1
Ghost3
Ghost4
Ghost5
Anther12
Embryo3
Embryo4
Embryo5
Style2
Root
Root_P
Leaf
Ovule1
Wall1
Anther11
Anther9
Anther10
Seed2
Anther78
Carpel78
Carpel9
Seedling
Wall5
Wall3
Wall4
Carpel12
Wall2

Figure 15 (previous page). Expression profiles of *FvWRKY* family members in different tissues, developmental stages and growth conditions of *F. vesca*. Color scale represents the expression level as log transformed TPM (Transcripts per Million) values. Tissue descriptions were taken from Hollender et al. (2012) and Li et al. (2019). SAM (shoot apical meristem), FM (floral meristem), REM (receptacle meristem), Anther7-8 (identified by stomium development and appearance of a preliminary lobed structure), Anther9 (microspore mother cells start meiosis), Anther10 (microspores are loose in the locule after callose wall holding tetrads disaggregates), Anther11 (pollen mitotic division occurs), anther12 (no visible change in anther development), Carpel7/8 (round carpel primordial reach the receptacle apex), Carpel9 (bowling pin shaped carpel primordial), Carpel10 (carpel is dived in almost equal apical and basal part by a central constriction), Carpel11 (style is elongated and became twice in length than the ovary base), Carpel12 (carpels have music note shape and styles are separated from each other), Cortex1 and Pith1 (flower just opened), Cortex2 and Pith2 (at about 3 DPA, when pollination occurs), Cortex3 and Pith3 (at about 6 DPA), Cortex4 and Pith4 (at about 9 DPA), Cortex5 and Pith5 (at about 12 DPA), Rg15D and Rg22D (Ruegen F7–4 receptacle tissue at 15 DPA and at 22 DPA, corresponding to green and white-turning stages respectively), Yw15D and Yw22D (Yellow Wonder 5AF7 receptacle tissue at 15 DPA and at 22 DPA, corresponding to green and white-turning stages respectively), Embryo3 and Ghost3 (embryo and seed without embryo inside at about 6 DPA characterized by heart shape), Embryo4 and Ghost4 (at about 9DPA, with immature cotyledons), Embryo5 and Ghost5 (ad about 12 DPA, mature embryo which fill up entire ovules), Leaf (young trifoliate leaves), Ovule1 and Pollen (collected from just open flower), Seed2 (complete achene from mature fruit), Seedling (complete seedling at 10 days post germination), Style1 (style and stigma from just open flowers), Style2 (style from flower at about 3 DPA), Wall1 (carpel wall from just open flower), Wall2 (carpel wall at about 3 DPA), Wall3 (carpel wall at about 6 DPA), Wall4 (carpel wall at about 9 DPA), Wall5 (carpel wall at about 12 DPA), Root (collected from 7 week old plants grown in aerated hydroponic culture) and Root_P (after 2 days of inoculation with *Phytophthora cactorum*). DPA: days post-anthesis.

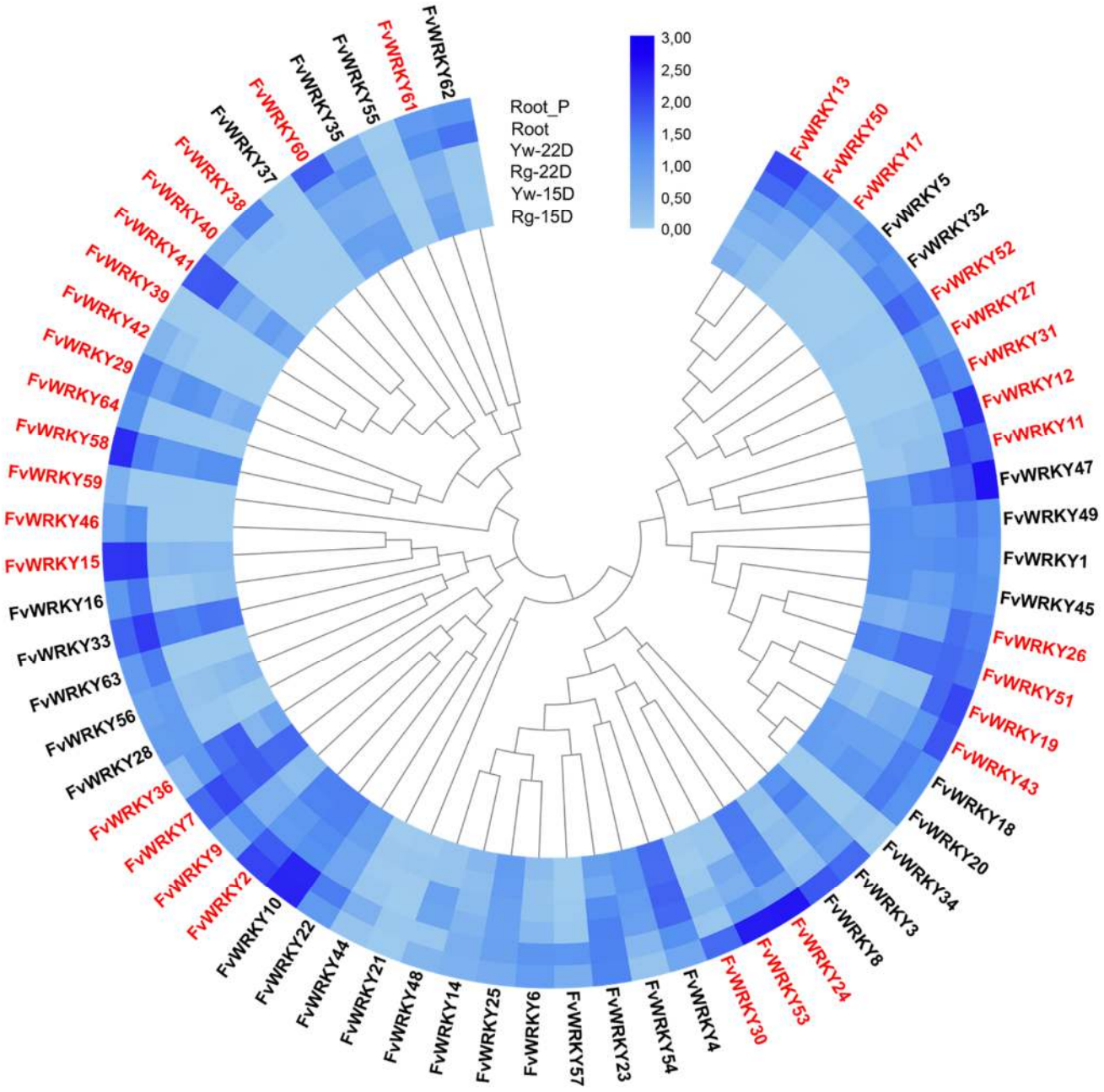


Figure 16. Phylogenetic tree of the FvWRKY family showing expression values during fruit ripening (Ruegen receptacle tissue at 15 DPA and at 22 DPA and Yellow Wonder receptacle tissue at 15 DPA and at 22 DPA) and roots (Root, collected from 7 week old plants grown in aerated hydroponic culture, and Root_P, after 2 days of inoculation with *Phytophthora cactorum*). FvWRKY paralogs are shown in red (see Table 6). Color scale represents the expression level as log transformed TPM (Transcripts per Million) values.

2.4.6. Expression analysis of the *FaWRKY* gene family in strawberry tissues, the fruit ripening process and defense responses

The recent publication of a high-quality annotated octoploid strawberry genome, and the availability of public RNA-seq data, provides a valuable chance to investigate the expression patterns of the *FaWRKY* gene family in different tissues and growth conditions at the level of homoeologs and paralogs, in order to understand their functions in key strawberry processes, such as fruit ripening and defense responses against pathogens. Accordingly, we have investigated two publicly available strawberry datasets, featuring both tissue expression profiles and transcriptomic changes along four stages of achene and receptacle ripening (Sanchez-Sevilla et al., 2017), and anthracnose defense response of strawberry leaves infected with the hemibiotrophic fungus *C. fructicola* (Zhang et al., 2018b). Both raw and differential gene expression values are provided in Supplementary Table S7 and S8, respectively.

As expected, the expression pattern for most of the *FaWRKY* genes varied among the different strawberry tissues investigated, showing many cases of preferential or even tissue-specific expression (Figure 17 and Supplementary Tables S7, S8 and S9). For example, *FaWRKY17B/17C.1/17C.2/17D* were detected and differentially expressed exclusively in root, whereas the homoeolog *FaWRKY17A* was expressed root, leaf and green receptacle. On the other hand, *FaWRKY24A/24B/24D* among others, showed a higher transcript accumulation in root than in any other tissues. Accordingly, the strawberry *FaWRKY24* homoeologs were identified as syntenic orthologs of AtWRKY75, which acts as negative regulator of root development as well as positive regulator of Pi stress responses (Devaiah et al., 2007).

Fruit ripening is a critical developmental process in strawberry, and fruit size, color or aroma, among others, are main agronomical traits. However, little is known about the involvement of members of the strawberry WRKY gene family in the genetic control of ripening and the major changes undergone in strawberry fruit during this important development process. Nevertheless, transcriptional reprogramming of *FaWRKY* genes is evident in both receptacle and achene from immature (green) to mature (red) fruit (Figure 18; Supplemetary Table S10 and Table S11). We have found that 74 *FaWRKYs* were significantly up- or down-regulated in receptacles during specific stages of ripening, while 102 were in their respective achenes (Figure 19A). Of these, 24 and 52 *FaWRKY* genes were differentially expressed only in receptacles or achenes, respectively (Figure 19B, Table 12). It is assumed that the hormonal balance between abscisic acid (ABA) and auxins synthetized in the achenes is primarily responsible, but not the only, for the changes leading to ripening in this non-climacteric fruit (Symons et al., 2012). So, it is not unreasonable to think that changes in the expression pattern of *FaWRKYs* might be controlled by this hormonal equilibrium. Supporting this, the expression in the receptacle of three strawberry *WRKYs (*Fv *gene28720, gene19478 and gene01340)* have previously been described to be activated by ABA and repressed by auxins in the receptacle, while another three WRKY genes *(*Fv *gene07210, gene03411 and gene09147*) were repressed by auxins and not affected by ABA

(Medina-Puche et al., 2016). Our transcriptomic analysis supports potential roles for five of these genes at some points in the course of the strawberry ripening. Thus, we have found that the homoeologous set of genes *FaWRKY48B/48C and FaWRKY53A/53B,* orthologs of the Fv *gene28720* and *gene01340* respectively, were differentially up-regulated in both receptacle and achene, and *FaWRKY48A* only in achene. On the other hand, *FaWRKY24D* (Fv *gene07210*) was differentially up-regulated in both receptacle and achene, but its homoeologous *FaWRKY24A* was down-regulated in receptacle and up-regulated in achene, from white to red stage. Also, the *gene03411* ortholog *FaWRKY17A* was up-regulated in red receptacle, whereas *gene09147* orthologs *FaWRKY9B/9C/9D* were down-regulated in achene and *FaWRKY9D* was up-regulated in receptacle. However, the *FaWRKY13* homoeologs, orthologs to the Fv *gene19478*, were not differentially expressed.

Besides, this analysis revealed that many paralogous and homoeologous genes were expressed differently in both fruit tissues (Table 12), suggesting that they may have undergone functional divergences in the strawberry fruit ripening. For example, *FaWRKY3C/3D, FaWRKY10A, FaWRKY15A/15B, FaWRKY20B, FaWRKY28A*, and *FaWRKY41D* were differentially expressed only in receptacle, whereas *FaWRKY3A*, *FaWRKY10C*, *FaWRKY15C/15D, FaWRKY20C, FaWRKY27D, and FaWRKY28C/28D* were only in achene tissue. Also, *FaWRKY2B.1/2D.1, FaWRKY6A, FaWRKY12A, FaWRKY16C/16D, FaWRKY33D, and FaWRKY42D* were expressed only in receptacle but *FaWRKY2A, FaWRKY6D, FaWRKY12C, FaWRKY16A, FaWRKY33A, and FaWRKY42B* were expressed in both receptacle and achene tissues. Other genes as *FaWRKY9B/9C, FaWRKY11B, FaWRKY22A, FaWRKY23B, FaWRKY26B.1/26D, FaWRKY48A, FaWRKY51C, FaWRKY55A/55D, FaWRKY56B/56D.1/56D.2* and *FaWRKY60D* were expressed in achenes, but *FaWRKY9D, FaWRKY11A, FaWRKY22B/22C/22D, FaWRKY23D, FaWRKY26A/26C, FaWRKY48B/48C, FaWRKY51A.1/51A.2, FaWRKY55C, FaWRKY56A.2/56C* and *FaWRKY60C* in both receptacle and achene tissues. In addition, *FaWRKY58A.2*, expressed only in receptacle whereas *FaWRKY58C* was expressed only in achenes and *FaWRKY58A.1/58B/58D.1/58D.2* in both fruit tissues. On the contrary, other *FaWRKY* homoeologous genes showed similar tissue specific expression pattern, suggesting that they may wield complementary roles and additive effects on the genetic regulation of fruit ripening. This is the case for *FaWRKY54A.1/54A.2/54B/54C,* expressed in receptacle; *FaWRKY1A/1B, FaWRKY7B/7C/7D, FaWRKY14A/14C.2, FaWRKY27C/27D, FaWRKY30A/30B, FaWRKY34A/34C, FaWRKY57A.2/57B/57C/57D.1/57D.2, FaWRKY63B.1/63D.1* or *FaWRKY64C/64D.1/64D.2,* expressed in achenes; and *FaWRKY19A/19B/19D, FaWRKY24A/24D, FaWRKY29A/29C/29D.1/29D.2, FaWRKY43A/43B.2/43D, FaWRKY45B/45C, FaWRKY47B/47C.1/47D, FaWRKY50A/50B/50C*, *FaWRKY53A/53B,* upregulated in both fruit tissues.

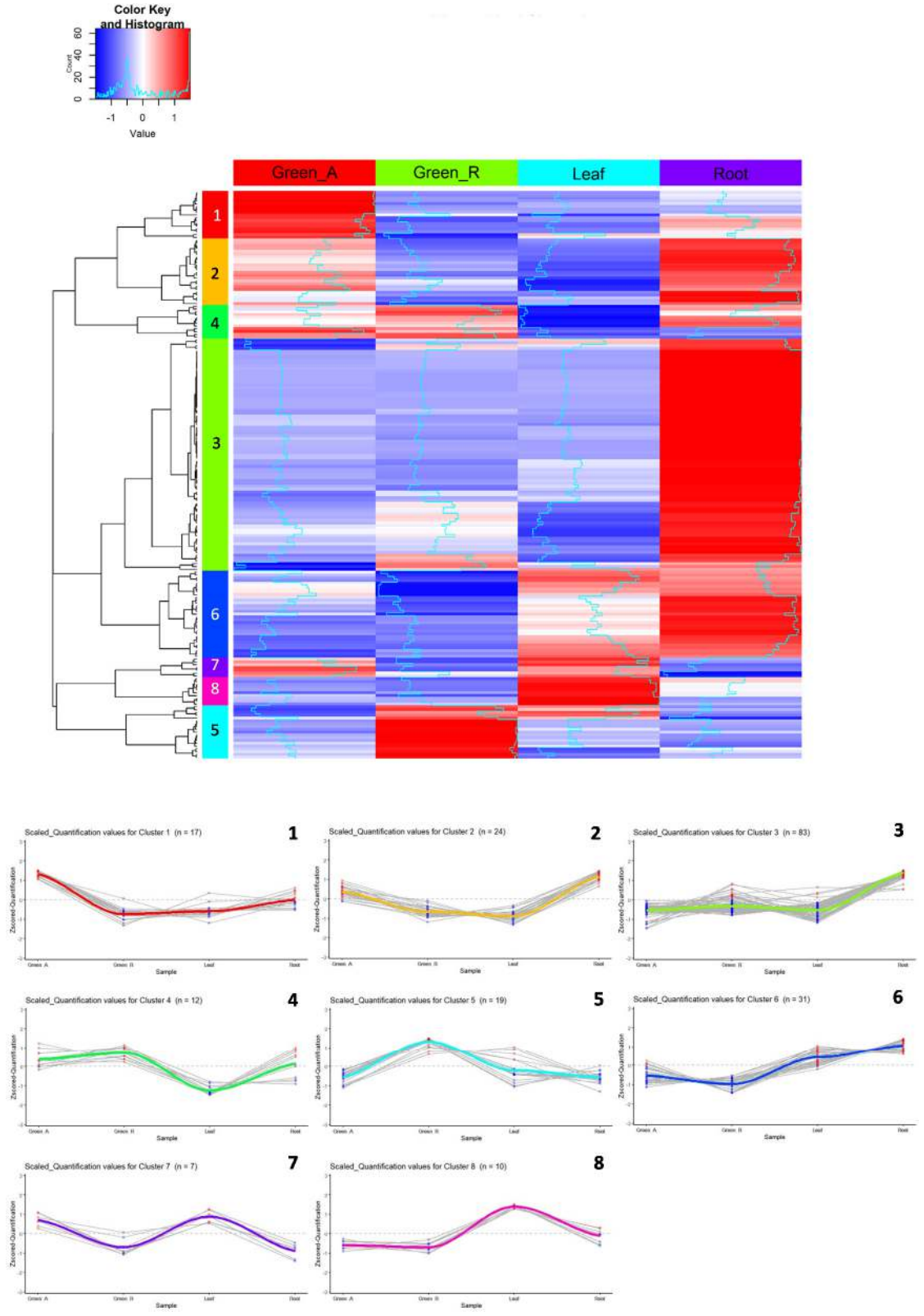


Figure 17. Expression and hierarchical clustering of *FaWRKY* genes in strawberry achene and receptacle of green fruits (Green_A and Green_R, respectively), leaves and root tissues. Heatmap represents gene expression ranged in a color scale from lowest (blue) to highest (red). Expression profiles of the statistically different groups (clusters) were depicted using same colors and numbers as in heatmap.

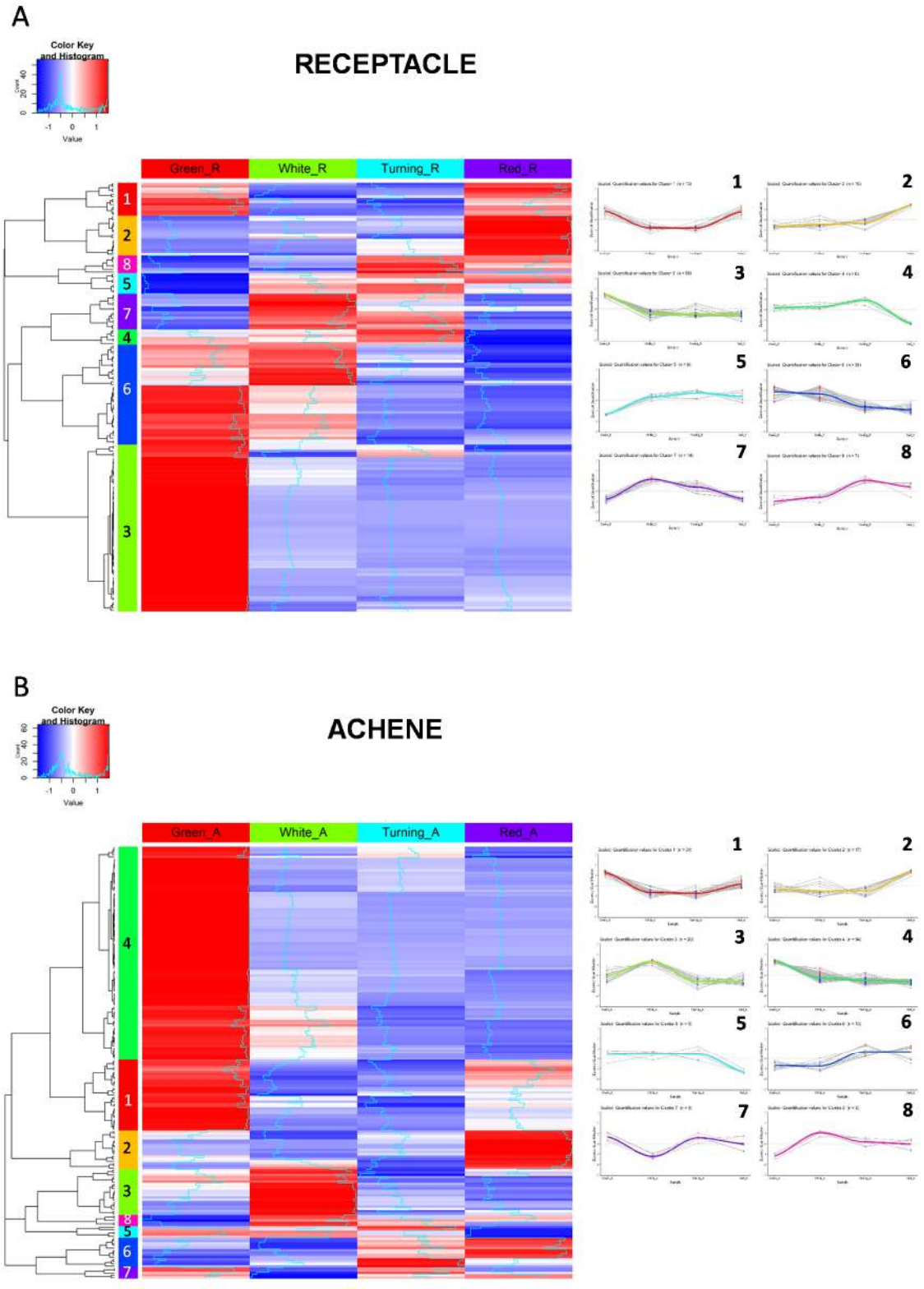


Figure 18. Expression and hierarchical clustering of *FaWRKY* genes during strawberry fruit ripening in receptacle (**A**) and achene (**B**) tissues. Heatmap represents gene expression ranged in a color scale from lowest (blue) to highest (red). Expression profiles of the statistically different (clusters) were depicted using same colors and numbers as in heatmap.

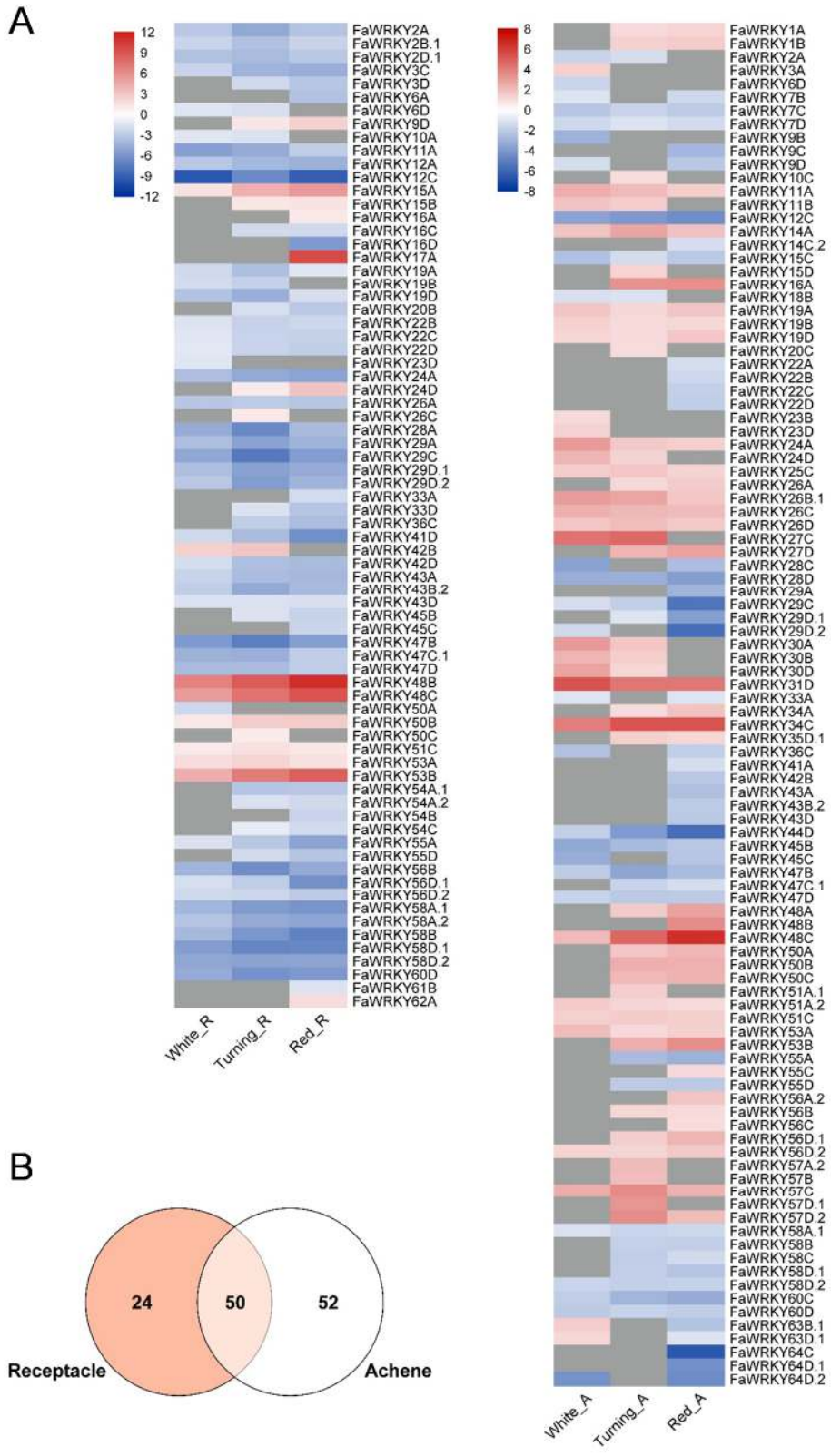


Figure 19. A, Heatmap of differentially expressed *FaWRKY* genes during strawberry fruit ripening in receptacle (left) and achene (right) tissues. Changes in gene expression, respect to green fruit tissues, are represented as log2 FC (fold change) if absolute values were >1 (padj<0.01), otherwise they were colored in grey. **B**, Venn diagram showing those genes which were differentially expressed in receptacle and achene only, or in both.

Table 12. Lists of differentially expressed *FaWRKYs* in fruit tissues during any stage of ripening. Label are set of homeologous and paralogous genes differentially expressed in receptacle and aquenes (*), in receptacle or in both tissues (1), in achene and both tissues (2), receptacle, achenes and both tissues (3).

Receptacle only	Achene only		Receptacle and achene	
FaWRKY2B.1 [1]	*FaWRKY1A*	*FaWRKY31D*	*FaWRKY2A* [1]	*FaWRKY45B*
FaWRKY2D.1 [1]	*FaWRKY1B*	*FaWRKY34A*	*FaWRKY6D* [1]	*FaWRKY45C*
*FaWRKY3C**	*FaWRKY3A**	*FaWRKY34C*	*FaWRKY9D* [2]	*FaWRKY47B*
*FaWRKY3D**	*FaWRKY7B*	*FaWRKY35D.1*	*FaWRKY11A* [2]	*FaWRKY47C.1*
FaWRKY6A [1]	*FaWRKY7C*	*FaWRKY41A**	*FaWRKY12C* [1]	*FaWRKY47D*
*FaWRKY10A**	*FaWRKY7D*	*FaWRKY44D*	*FaWRKY16A* [1]	*FaWRKY48B* [2]
FaWRKY12A [1]	*FaWRKY9B* [2]	*FaWRKY48A* [2]	*FaWRKY19A*	*FaWRKY48C* [2]
*FaWRKY15A**	*FaWRKY9C* [2]	*FaWRKY51A.1* [2]	*FaWRKY19B*	*FaWRKY50A*
*FaWRKY15B**	*FaWRKY10C**	*FaWRKY51A.2* [2]	*FaWRKY19D*	*FaWRKY50B*
FaWRKY16C [1]	*FaWRKY11B* [2]	*FaWRKY55C* [2]	*FaWRKY22B* [2]	*FaWRKY50C*
FaWRKY16D [1]	*FaWRKY14A*	*FaWRKY56A.2* [2]	*FaWRKY22C* [2]	*FaWRKY51C* [2]
FaWRKY17A	*FaWRKY14C.2*	*FaWRKY56C* [2]	*FaWRKY22D* [2]	*FaWRKY53A*
*FaWRKY20B**	*FaWRKY15C**	*FaWRKY57A.2*	*FaWRKY23D* [2]	*FaWRKY53B*
*FaWRKY28A**	*FaWRKY15D**	*FaWRKY57B*	*FaWRKY24A*	*FaWRKY55A* [2]
FaWRKY33D [1]	*FaWRKY18B*	*FaWRKY57C*	*FaWRKY24D*	*FaWRKY55D* [2]
*FaWRKY41D**	*FaWRKY20C**	*FaWRKY57D.1*	*FaWRKY26A* [2]	*FaWRKY56B* [2]
FaWRKY42D [1]	*FaWRKY22A* [2]	*FaWRKY57D.2*	*FaWRKY26C* [2]	*FaWRKY56D.1* [2]
FaWRKY54A.1	*FaWRKY23B* [2]	*FaWRKY58C* [3]	*FaWRKY29A*	*FaWRKY56D.2* [2]
FaWRKY54A.2	*FaWRKY25C*	*FaWRKY60C* [2]	*FaWRKY29C*	*FaWRKY58A.1* [3]
FaWRKY54B	*FaWRKY26B.1* [2]	*FaWRKY63B.1*	*FaWRKY29D.1*	*FaWRKY58B* [3]
FaWRKY54C	*FaWRKY26D* [2]	*FaWRKY63D.1*	*FaWRKY29D.2*	*FaWRKY58D.1* [3]
FaWRKY58A.2 [3]	*FaWRKY27C*	*FaWRKY64C*	*FaWRKY33A* [1]	*FaWRKY58D.2* [3]
FaWRKY61B	*FaWRKY27D*	*FaWRKY64D.1*	*FaWRKY36C*	*FaWRKY60D* [2]
FaWRKY62A	*FaWRKY28C**	*FaWRKY64D.2*	*FaWRKY42B* [1]	
	*FaWRKY28D**		*FaWRKY43A*	
	FaWRKY30A		*FaWRKY43B.2*	
	FaWRKY30B		*FaWRKY43D*	
	FaWRKY30D			

It is known that WRKY TFs have important roles in defense responses against pathogenic fungi (Chen et al., 2019). The expression patterns of several *FaWRKY* genes fluctuated in strawberry leaves challenged with *C. fructicola* (Figure 20 and Supplementary Table S12). A total of 53 *FaWRKYs* were differentially expressed, 41 up-regulated and 12 downregulated (Figure 21), most of which were homologs of well-known defense-related WRKYs. Among the *FaWRKY* upregulated genes with higher expression and having known functions in plant defense, we found WRKY8/28 probable orthologs *FaWRKY14A, FaWRKY14B, FaWRKY14C.2* and *FaWRKY14D*; WRKY54/70 orthologs *FaWRKY38B*, *FaWRKY39A* and *FaWRKY60B*; WRKY75 orthologs *FaWRKY24A* and *FaWRKY24D*; WRKY50/51 orthologs *FaWRKY3C, FaWRKY8A, FaWRKY8B, FaWRKY8C* and *FaWRKY8D*; WRKY72 ortholog *FaWRKY52A*; and WRKY18 ortholog *FaWRKY11A*. On the other hand, among the downregulated *FaWRKY* genes were the WRKY53 orthologs *FaWRKY29A, FaWRKY29B, FaWRKY29D.*1 and *FaWRKY29D.2.*

Interestingly, several of the above mentioned genes are involved in negative regulation of JA-mediated, but positive regulation of SA-mediated defense responses (Li et al., 2004;Li et al., 2006;Bhattarai et al., 2010;Gao et al., 2011;van Verk et al., 2011;Hu et al., 2012), while WRKY18 regulates positively the expression of some key JA-signalling genes in *Arabidopsis* (Pandey et al., 2010;Birkenbihl et al., 2017). It is worth to mention that AtWRKY33 high-homology sharing genes *FaWRKY43B.2* and *FaWRKY43D* are downregulated along the time course of infection. *Arabidopsis* WRKY33 is a key positive regulator of the JA-mediated defense response against necrotrophic fungi, also showing antagonistic effect with the SA-mediated pathway (Zheng et al., 2006) and AtWRKY33 orthologs seem to share similar functions in other plant species, including the woodland strawberry (Wei et al., 2018). Although this downregulation was already noticed by Zhang et al. (Zhang et al., 2018b), it contrasts markedly with the results reported in other strawberry plant tissues of either resistant or susceptible cultivars challenged with *C. gloeosporioides* (Wang et al., 2017) or *C. acutatum* (Amil-Ruiz et al., 2016;Garrido-Gala et al., 2019). However, it could be explained if subtle differences in the pathogen-specific strategies do exist among *Colletotrichum* species to manipulate the host defense in different strawberry tissues.

Homoeologs of *FaWRKY24* and *FaWRKY53* are both AtWRKY75-like factors and as mentioned before, FaWRKY24 (formely, FaWRKY1) and AtWRKY75 have been previously reported to play important roles as positive regulators of plant defense in *A. thaliana* against *P. syringae* (Encinas-Villarejo et al., 2009; Amil-Ruiz et al., 2016). Both, the Fv paralogs *FvWRKY24* and *FvWRKY53* and their Fa orthologs *FaWRKY24A/24D* and *FaWRKY53B/53C/53D* were also detected to be up-regulated in response to different pathogens (see Figures 15 and 21). In addition, a recent research has shown that WRKY75 is a positive regulator of the JA-mediated defense to the necrotrophic fungi *B. cinerea* and *A. brassicicola* (Chen et al., 2021). Interestingly, *FaWRKY24* (*FaWRKY1*) homoeologs are also up-regulated during strawberry fruit ripening, as well as they have been reported as negative regulators of resistance to *C. acutatum* in strawberry fruit (Higuera et al., 2019). This illustrates the dual roles that the strawberry

AtWRKY75-like factors may play against different pathogens or in different plant tissues.

Among the R protein-WRKY subclass, *FaWRKY55D* was the only one differentially expressed at early stages of the infection. Thus, it putatively should be included among the repertory of strawberry *WRKY* genes responsive to pathogenic fungi. Several homoeologs of the other *FaRWRKY* genes (*FaWRKY35, FaWRKY61* and *FaWRKY62*) were expressed in both mock-treated and infected tissues. However, their transcript abundances did not change significantly in the course of infection. Most of them showed relatively high transcript abundances also in leaves, roots and fruit tissues, suggesting that changes in their gene expression, if any, are either not needed to play positive roles against pathogens or these changes are pathogen-specific deployed or involved in other plant processes. Nevertheless, their roles in the strawberry physiology remains unknown, as the vast majority of other R protein-WRKY in several species.

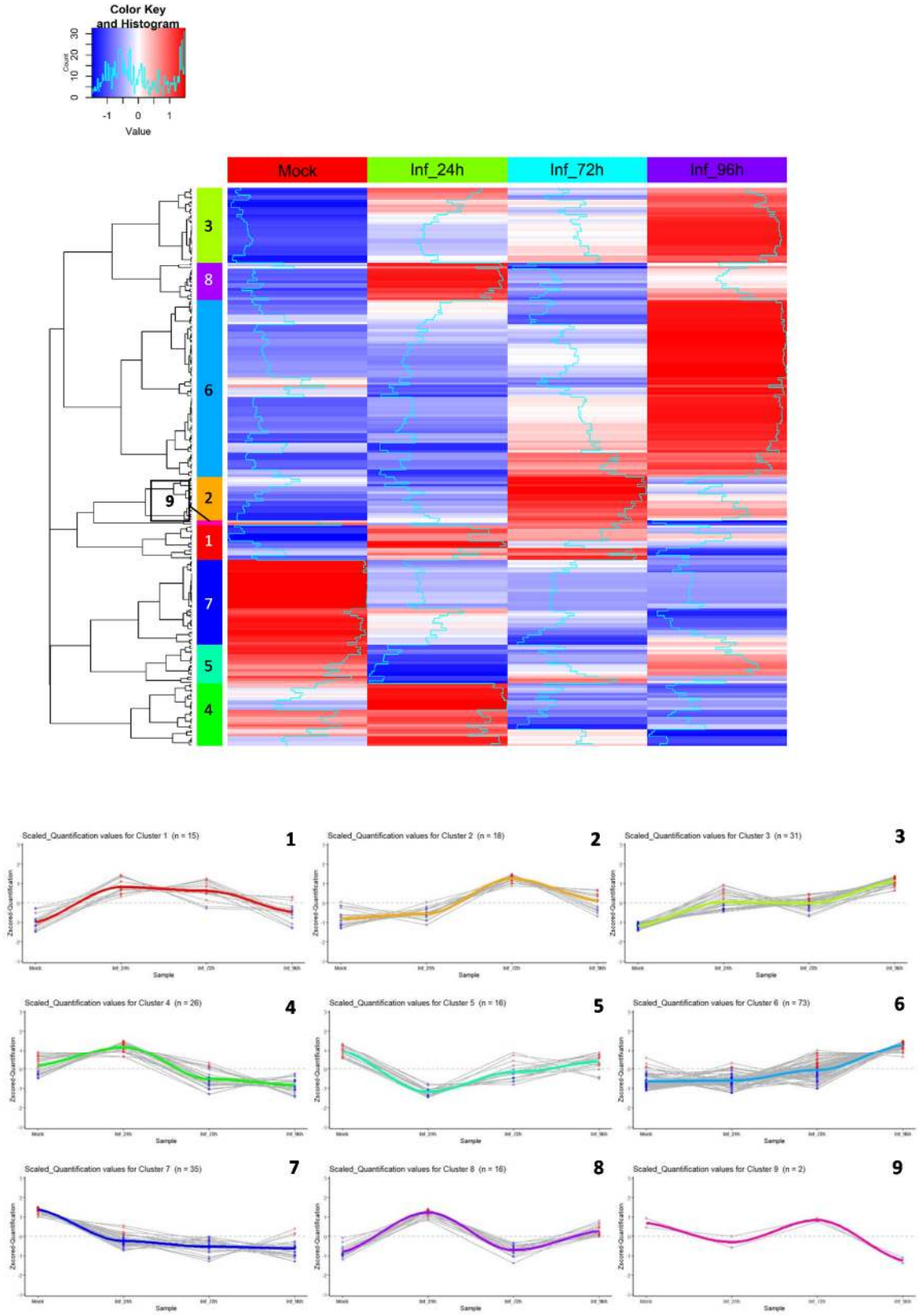


Figure 20. Expression and hierarchical clustering of *FaWRKY* genes in strawberry leaves, inoculated with mock or *C. fructicola* spores and collected at 24, 72 and 96 hours. Heatmap represents gene expression ranged in a color scale from lowest (blue) to highest (red). Expression profiles of the statistically different groups (clusters) were depicted using same colors and numbers as in heatmap.

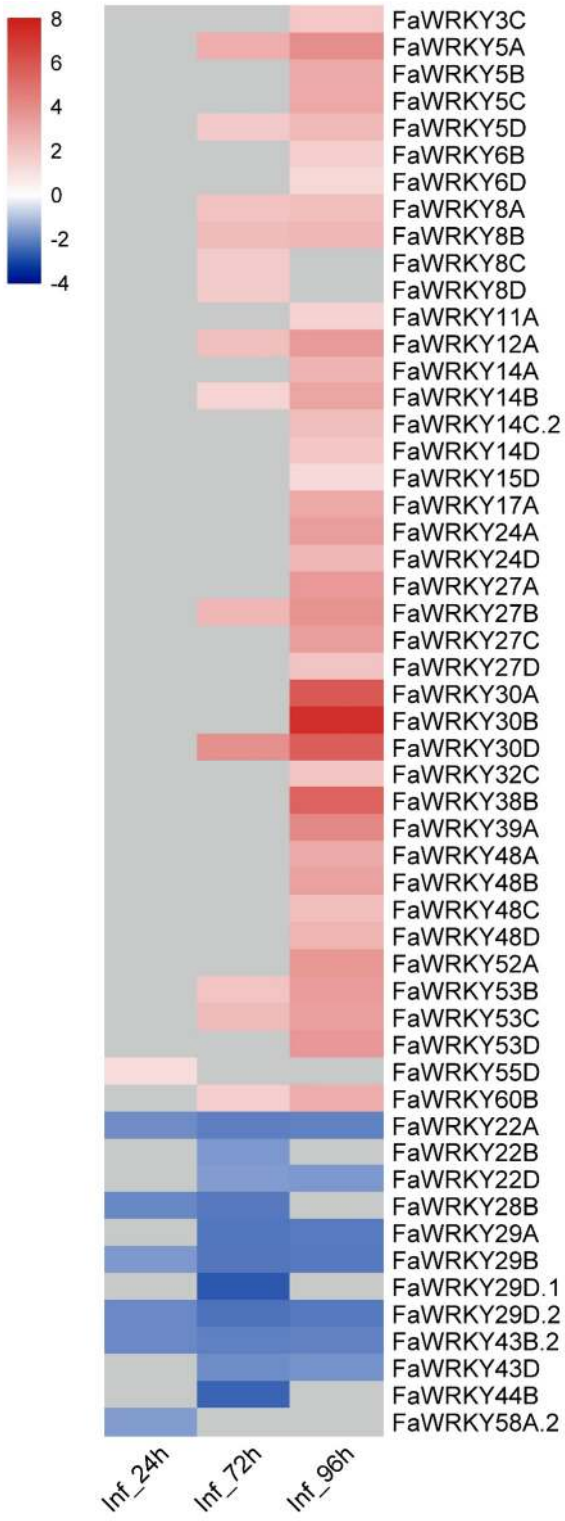


Figure 21. Heatmap of differentially expressed *FaWRKY* genes in response to *C. fructicola* infection in leaves. Changes in gene expression, respect to mock treated control leaves, are represented as log2 FC (fold change) if absolute values were >1 (padj<0.01), in a color scale from lowest (blue) to highest (red), otherwise they were colored in grey.

2.5. Conclusions

The present study represents a comprehensive update of the strawberry WRKY family by using the most complete ultimate and accurate genome assemblies and gene annotations available in strawberry. Using the most up-to-date data has allowed us to perform highly precise analyses and characterization of the WRKY gene family composition, its evolutionary history and gene expression pattern in strawberry, particularly in the cultivated hybrid *Fragaria x ananassa*. Importantly, we have described for the first time the WRKY homoeologs and gene duplications found in the octoploid.

Thus, a total of 64 WRKY genes are present in the genome of the diploid *Fragaria vesca* and 257 corresponding WRKY orthologs in the genome of the cultivated allo-octoploid *Fragaria x ananassa* (cv. Camarosa). Our results show that the strawberry WRKY family preserve a high degree of conservation between both wild type and cultivated species. Synteny analysis showed that *FaWRKY* genes are largely syntenic and collinear with their *FvWRKY* orthologs, despite some gene loses and synteny breaks detected in the octoploid, probably as a result of genomic rearrangements derived from hybridization and polyploidy. The strawberry WRKY family has been expanded by ancient WGD events, originating several segmental and tandem duplications resulting in several paralogous genes. Moreover, most *FaWRKY* paralogs are not mirrored in *F. vesca*, and could be inherited from the parental octoploid species *F. virginiana* and *F. chiloensis*, in which polyploidization could originate synteny breaks among ancestral paralogs as well as new gene duplications.

Gene expression is related with functionality, and its analysis can shed light about the biological roles of the genes studied. Thus, the expression patterns of most WRKY genes differed in the strawberry tissues here considered, evidencing differences in functional relevance across different tissues and growth conditions. Importantly, many strawberry WRKYs changed remarkably their transcript level or were differentially expressed in defense responses and during fruit ripening stages, indicating their importance in these key biological processes. In addition, differences in the expression pattern detected among several *FaWRKY* paralogs and homoeologs, point out either a finely regulated gene expression strategy to achieve putative additive genetic effects or evolutionary functional divergences, These data can help future studies to deepen our understanding of the strawberry WRKY TF regulatory roles as well as be considered in breeding programs.

2.6. References

Afgan, E., Baker, D., Van Den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., et al. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res* 44**,** W3-W10.doi: 10.1093/nar/gkw343

Alves, M.S., Dadalto, S.P., Goncalves, A.B., De Souza, G.B., Barros, V.A., and Fietto, L.G. (2014). Transcription Factor Functional Protein-Protein Interactions in Plant Defense Responses. *Proteomes* 2**,** 85-106.doi: 10.3390/proteomes2010085

Amborella Genome, P. (2013). The Amborella genome and the evolution of flowering plants. *Science* 342**,** 1241089.doi: 10.1126/science.1241089

Amil-Ruiz, F., Garrido-Gala, J., Blanco-Portales, R., Folta, K.M., Munoz-Blanco, J., and Caballero, J.L. (2013). Identification and validation of reference genes for transcript normalization in strawberry (Fragaria x ananassa) defense responses. *PLoS One* 8**,** e70603.doi: 10.1371/journal.pone.0070603

Amil-Ruiz, F., Garrido-Gala, J., Gadea, J., Blanco-Portales, R., Munoz-Merida, A., Trelles, O., et al. (2016). Partial Activation of SA- and JA-Defensive Pathways in Strawberry upon Colletotrichum acutatum Interaction. *Front Plant Sci* 7**,** 1036.doi: 10.3389/fpls.2016.01036

Bailey, P.C., Schudoma, C., Jackson, W., Baggs, E., Dagdas, G., Haerty, W., et al. (2018). Dominant integration locus drives continuous diversification of plant immune receptors with exogenous domain fusions. *Genome Biol* 19**,** 23.doi: 10.1186/s13059-018-1392-6

Bakshi, M., and Oelmuller, R. (2014). WRKY transcription factors: Jack of many trades in plants. *Plant Signal Behav* 9**,** e27700.doi: 10.4161/psb.27700

Bashir, T., Chandra Mishra, R., Hasan, M.M., Mohanta, T.K., and Bae, H. (2018). Effect of Hybridization on Somatic Mutations and Genomic Rearrangements in Plants. *Int J Mol Sci* 19.doi: 10.3390/ijms19123758

Besseau, S., Li, J., and Palva, E.T. (2012). WRKY54 and WRKY70 co-operate as negative regulators of leaf senescence in Arabidopsis thaliana. *J Exp Bot* 63**,** 2667-2679.doi: 10.1093/jxb/err450

Bhattarai, K.K., Atamian, H.S., Kaloshian, I., and Eulgem, T. (2010). WRKY72-type transcription factors contribute to basal immunity in tomato and Arabidopsis as well as gene-for-gene resistance mediated by the tomato R gene Mi-1. *Plant J* 63**,** 229-240.doi: 10.1111/j.1365-313X.2010.04232.x

Birkenbihl, R.P., Kracher, B., Roccaro, M., and Somssich, I.E. (2017). Induced Genome-Wide Binding of Three Arabidopsis WRKY Transcription Factors during Early MAMP-Triggered Immunity. *Plant Cell* 29**,** 20-38.doi: 10.1105/tpc.16.00681

Birkenbihl, R.P., Kracher, B., Ross, A., Kramer, K., Finkemeier, I., and Somssich, I.E. (2018). Principles and characteristics of the Arabidopsis WRKY regulatory network during early MAMP-triggered immunity. *Plant J* 96**,** 487-502.doi: 10.1111/tpj.14043

Chanderbali, A.S., Berger, B.A., Howarth, D.G., Soltis, D.E., and Soltis, P.S. (2017). Evolution of floral diversity: genomics, genes and gamma. *Philos Trans R Soc Lond B Biol Sci* 372.doi: 10.1098/rstb.2015.0509

Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y., et al. (2020). TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol Plant* 13**,** 1194-1202.doi: 10.1016/j.molp.2020.06.009

Chen, J., Nolan, T.M., Ye, H., Zhang, M., Tong, H., Xin, P., et al. (2017a). Arabidopsis WRKY46, WRKY54, and WRKY70 Transcription Factors Are Involved in Brassinosteroid-Regulated Plant Growth and Drought Responses. *Plant Cell* 29**,** 1425-1439.doi: 10.1105/tpc.17.00364

Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., and Yu, D. (2012). The role of WRKY transcription factors in plant abiotic stresses. *Biochim Biophys Acta* 1819**,** 120-128.doi: 10.1016/j.bbagrm.2011.09.002

Chen, L., Xiang, S., Chen, Y., Li, D., and Yu, D. (2017b). Arabidopsis WRKY45 Interacts with the DELLA Protein RGL1 to Positively Regulate Age-Triggered Leaf Senescence. *Mol Plant* 10**,** 1174-1189.doi: 10.1016/j.molp.2017.07.008

Chen, L., Zhang, L., Xiang, S., Chen, Y., Zhang, H., and Yu, D. (2021). The transcription factor WRKY75 positively regulates jasmonate-mediated plant defense to necrotrophic fungal pathogens. *J Exp Bot* 72**,** 1473-1489.doi: 10.1093/jxb/eraa529

Chen, X., Li, C., Wang, H., and Guo, Z. (2019). WRKY transcription factors: evolution, binding, and action. *Phytopathology Research* 1.doi: 10.1186/s42483-019-0022-x

Cheng, X., Zhao, Y., Jiang, Q., Yang, J., Zhao, W., Taylor, I.A., et al. (2019). Structural basis of dimerization and dual W-box DNA recognition by rice WRKY domain. *Nucleic Acids Res* 47**,** 4308-4318.doi: 10.1093/nar/gkz113

Ciolkowski, I., Wanke, D., Birkenbihl, R.P., and Somssich, I.E. (2008). Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant Mol Biol* 68**,** 81-92.doi: 10.1007/s11103-008-9353-1

Cooper, G.M., and Brown, C.D. (2008). Qualifying the relationship between sequence conservation and molecular function. *Genome Res* 18**,** 201-205.doi: 10.1101/gr.7205808

Darwish, O., Shahan, R., Liu, Z., Slovin, J.P., and Alkharouf, N.W. (2015). Re-annotation of the woodland strawberry (Fragaria vesca) genome. *BMC Genomics* 16**,** 29.doi: 10.1186/s12864-015-1221-1

De La Torre, A.R., Li, Z., Van De Peer, Y., and Ingvarsson, P.K. (2017). Contrasting Rates of Molecular Evolution and Patterns of Selection among Gymnosperms and Flowering Plants. *Mol Biol Evol* 34**,** 1363-1377.doi: 10.1093/molbev/msx069

Deslandes, L., Olivier, J., Peeters, N., Feng, D.X., Khounlotham, M., Boucher, C., et al. (2003). Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proc Natl Acad Sci U S A* 100**,** 8024-8029.doi: 10.1073/pnas.1230660100

Devaiah, B.N., Karthikeyan, A.S., and Raghothama, K.G. (2007). WRKY75 transcription factor is a modulator of phosphate acquisition and root development in Arabidopsis. *Plant Physiol* 143**,** 1789-1801.doi: 10.1104/pp.106.093971

Dimeglio, L.M., Staudt, G., Yu, H., and Davis, T.M. (2014). A phylogenetic analysis of the genus Fragaria (strawberry) using intron-containing sequence from the ADH-1 gene. *PLoS One* 9**,** e102237.doi: 10.1371/journal.pone.0102237

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29**,** 15-21.doi: 10.1093/bioinformatics/bts635

Doll, J., Muth, M., Riester, L., Nebel, S., Bresson, J., Lee, H.C., et al. (2019). Arabidopsis thaliana WRKY25 Transcription Factor Mediates Oxidative Stress Tolerance and Regulates Senescence in a Redox-Dependent Manner. *Front Plant Sci* 10**,** 1734.doi: 10.3389/fpls.2019.01734

Edger, P.P., Mckain, M.R., Yocca, A.E., Knapp, S.J., Qiao, Q., and Zhang, T. (2020). Reply to: Revisiting the origin of octoploid strawberry. *Nat Genet* 52**,** 5-7.doi: 10.1038/s41588-019-0544-2

Edger, P.P., Poorten, T.J., Vanburen, R., Hardigan, M.A., Colle, M., Mckain, M.R., et al. (2019). Origin and evolution of the octoploid strawberry genome. *Nat Genet* 51**,** 541-547.doi: 10.1038/s41588-019-0356-4

Edger, P.P., Vanburen, R., Colle, M., Poorten, T.J., Wai, C.M., Niederhuth, C.E., et al. (2018). Single-molecule sequencing and optical mapping yields an improved genome of

woodland strawberry (Fragaria vesca) with chromosome-scale contiguity. *Gigascience* 7**,** 1-7.doi: 10.1093/gigascience/gix124

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., et al. (2019). The Pfam protein families database in 2019. *Nucleic Acids Res* 47**,** D427-D432.doi: 10.1093/nar/gky995

Emms, D.M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 16**,** 157.doi: 10.1186/s13059-015-0721-2

Encinas-Villarejo, S., Maldonado, A.M., Amil-Ruiz, F., De Los Santos, B., Romero, F., Pliego-Alfaro, F., et al. (2009). Evidence for a positive regulatory role of strawberry (Fragaria×ananassa) Fa WRKY1 and Arabidopsis At WRKY75 proteins in resistance. *Journal of Experimental Botany* 60**,** 3043-3065.doi: 10.1093/jxb/erp152

Eulgem, T. (2006). Dissecting the WRKY web of plant defense regulators. *PLoS Pathog* 2**,** e126.doi: 10.1371/journal.ppat.0020126

Eulgem, T., Rushton, P.J., Robatzek, S., and Somssich, I.E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci* 5**,** 199-206.doi: 10.1016/s1360-1385(00)01600-9

Feng, C., Wang, J., Harris, A.J., Folta, K.M., Zhao, M., and Kang, M. (2020). Tracing the diploid ancestry of the cultivated octoploid strawberry. *Mol Biol Evol*.doi: 10.1093/molbev/msaa238

Gao, Q.M., Venugopal, S., Navarre, D., and Kachroo, A. (2011). Low oleic acid-derived repression of jasmonic acid-inducible defense responses requires the WRKY50 and WRKY51 proteins. *Plant Physiol* 155**,** 464-476.doi: 10.1104/pp.110.166876

Garrido-Gala, J., Higuera, J.J., Munoz-Blanco, J., Amil-Ruiz, F., and Caballero, J.L. (2019). The VQ motif-containing proteins in the diploid and octoploid strawberry. *Sci Rep* 9**,** 4942.doi: 10.1038/s41598-019-41210-4

Glover, N.M., Redestig, H., and Dessimoz, C. (2016). Homoeologs: What Are They and How Do We Infer Them? *Trends Plant Sci* 21**,** 609-621.doi: 10.1016/j.tplants.2016.02.005

Guan, Y., Meng, X., Khanna, R., Lamontagne, E., Liu, Y., and Zhang, S. (2014). Phosphorylation of a WRKY transcription factor by MAPKs is required for pollen development and function in Arabidopsis. *PLoS Genet* 10**,** e1004384.doi: 10.1371/journal.pgen.1004384

Guo, C., Guo, R., Xu, X., Gao, M., Li, X., Song, J., et al. (2014). Evolution and expression analysis of the grape (Vitis vinifera L.) WRKY gene family. *J Exp Bot* 65**,** 1513-1528.doi: 10.1093/jxb/eru007

Guo, H., Ahn, H.K., Sklenar, J., Huang, J., Ma, Y., Ding, P., et al. (2020). Phosphorylation-Regulated Activation of the Arabidopsis RRS1-R/RPS4 Immune Receptor Complex Reveals Two Distinct Effector Recognition Mechanisms. *Cell Host Microbe* 27**,** 769-781 e766.doi: 10.1016/j.chom.2020.03.008

Haas, B.J., Delcher, A.L., Wortman, J.R., and Salzberg, S.L. (2004). DAGchainer: a tool for mining segmental genome duplications and synteny. *Bioinformatics* 20**,** 3643-3646.doi: 10.1093/bioinformatics/bth397

Hafidh, S., Potesil, D., Muller, K., Fila, J., Michailidis, C., Herrmannova, A., et al. (2018). Dynamics of the Pollen Sequestrome Defined by Subcellular Coupled Omics. *Plant Physiol* 178**,** 258-282.doi: 10.1104/pp.18.00648

Haug-Baltzell, A., Stephens, S.A., Davey, S., Scheidegger, C.E., and Lyons, E. (2017). SynMap2 and SynMap3D: web-based whole-genome synteny browsers. *Bioinformatics* 33**,** 2197-2198.doi: 10.1093/bioinformatics/btx144

Higuera, J.J., Garrido-Gala, J., Lekhbou, A., Arjona-Girona, I., Amil-Ruiz, F., Mercado, J.A., et al. (2019). The Strawberry FaWRKY1 Transcription Factor Negatively Regulates Resistance to Colletotrichum acutatum in Fruit Upon Infection. *Front Plant Sci* 10**,** 480.doi: 10.3389/fpls.2019.00480

Hollender, C.A., Geretz, A.C., Slovin, J.P., and Liu, Z. (2012). Flower and early fruit development in a diploid strawberry, Fragaria vesca. *Planta* 235**,** 1123-1139.doi: 10.1007/s00425-011-1562-1

Hu, Y., Dong, Q., and Yu, D. (2012). Arabidopsis WRKY46 coordinates with WRKY70 and WRKY53 in basal resistance against pathogen Pseudomonas syringae. *Plant Sci* 185-186**,** 288-297.doi: 10.1016/j.plantsci.2011.12.003

Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., Von Mering, C., et al. (2017). Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Mol Biol Evol* 34**,** 2115-2122.doi: 10.1093/molbev/msx148

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernandez-Plaza, A., Forslund, S.K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 47**,** D309-D314.doi: 10.1093/nar/gky1085

Hughes, A.L., Friedman, R., Ekollu, V., and Rose, J.R. (2003). Non-random association of transposable elements with duplicated genomic blocks in Arabidopsis thaliana. *Mol Phylogenet Evol* 29**,** 410-416.doi: 10.1016/s1055-7903(03)00262-8

Hussain, R.M.F., Sheikh, A.H., Haider, I., Quareshy, M., and Linthorst, H.J.M. (2018). Arabidopsis WRKY50 and TGA Transcription Factors Synergistically Activate Expression of PR1. *Front Plant Sci* 9**,** 930.doi: 10.3389/fpls.2018.00930

Jiao, Y., Leebens-Mack, J., Ayyampalayam, S., Bowers, J.E., Mckain, M.R., Mcneal, J., et al. (2012). A genome triplication associated with early diversification of the core eudicots. *Genome Biol* 13**,** R3.doi: 10.1186/gb-2012-13-1-r3

Jin, J., Tian, F., Yang, D.C., Meng, Y.Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res* 45**,** D1040-D1045.doi: 10.1093/nar/gkw982

Jung, S., Lee, T., Cheng, C.H., Buble, K., Zheng, P., Yu, J., et al. (2019). 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. *Nucleic Acids Res* 47**,** D1137-D1145.doi: 10.1093/nar/gky1000

Kim, K.D., El Baidouri, M., Abernathy, B., Iwata-Otsubo, A., Chavarro, C., Gonzales, M., et al. (2015). A Comparative Epigenomic Analysis of Polyploidy-Derived Genes in Soybean and Common Bean. *Plant Physiol* 168**,** 1433-1447.doi: 10.1104/pp.15.00408

Koonin, E.V. (2005). Orthologs, paralogs, and evolutionary genomics. *Annu Rev Genet* 39**,** 309-338.doi: 10.1146/annurev.genet.39.073003.114725

Kopylova, E., Noe, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28**,** 3211-3217.doi: 10.1093/bioinformatics/bts611

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33**,** 1870-1874.doi: 10.1093/molbev/msw054

Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47**,** W256-W259.doi: 10.1093/nar/gkz239

Li, J., Brader, G., Kariola, T., and Palva, E.T. (2006). WRKY70 modulates the selection of signaling pathways in plant defense. *Plant J* 46**,** 477-491.doi: 10.1111/j.1365-313X.2006.02712.x

Li, J., Brader, G., and Palva, E.T. (2004). The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. *Plant Cell* 16**,** 319-331.doi: 10.1105/tpc.016980

Li, J., Zhong, R., and Palva, E.T. (2017). WRKY70 and its homolog WRKY54 negatively modulate the cell wall-associated defenses to necrotrophic pathogens in Arabidopsis. *PLoS One* 12**,** e0183731.doi: 10.1371/journal.pone.0183731

Li, Y., Pi, M., Gao, Q., Liu, Z., and Kang, C. (2019). Updated annotation of the wild strawberry Fragaria vesca V4 genome. *Hortic Res* 6**,** 61.doi: 10.1038/s41438-019-0142-6

Li, Y., Wei, W., Feng, J., Luo, H., Pi, M., Liu, Z., et al. (2018). Genome re-annotation of the wild strawberry Fragaria vesca using extensive Illumina- and SMRT-based RNA-seq datasets. *DNA Res* 25**,** 61-70.doi: 10.1093/dnares/dsx038

Liston, A., Wei, N., Tennessen, J.A., Li, J., Dong, M., and Ashman, T.L. (2020). Revisiting the origin of octoploid strawberry. *Nat Genet* 52**,** 2-4.doi: 10.1038/s41588-019-0543-3

Liu, D., Hunt, M., and Tsai, I.J. (2018). Inferring synteny between genome assemblies: a systematic evaluation. *BMC Bioinformatics* 19**,** 26.doi: 10.1186/s12859-018-2026-4

Luo, M.C., You, F.M., Li, P., Wang, J.R., Zhu, T., Dandekar, A.M., et al. (2015). Synteny analysis in Rosids with a walnut physical map reveals slow genome evolution in long-lived woody perennials. *BMC Genomics* 16**,** 707.doi: 10.1186/s12864-015-1906-5

Lyons, E., and Freeling, M. (2008). How to usefully compare homologous plant genes and chromosomes as DNA sequences. *Plant J* 53**,** 661-673.doi: 10.1111/j.1365-313X.2007.03326.x

Machens, F., Becker, M., Umrath, F., and Hehl, R. (2014). Identification of a novel type of WRKY transcription factor binding site in elicitor-responsive cis-sequences from Arabidopsis thaliana. *Plant Mol Biol* 84**,** 371-385.doi: 10.1007/s11103-013-0136-y

Maeo, K., Hayashi, S., Kojima-Suzuki, H., Morikami, A., and Nakamura, K. (2001). Role of conserved residues of the WRKY domain in the DNA-binding of tobacco WRKY family proteins. *Biosci Biotechnol Biochem* 65**,** 2428-2436.doi: 10.1271/bbb.65.2428

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., et al. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45**,** D200-D203.doi: 10.1093/nar/gkw1129

Medina-Puche, L., Blanco-Portales, R., Molina-Hidalgo, F.J., Cumplido-Laso, G., Garcia-Caparros, N., Moyano-Canete, E., et al. (2016). Extensive transcriptomic studies on the roles played by abscisic acid and auxins in the development and ripening of strawberry fruits. *Funct Integr Genomics* 16**,** 671-692.doi: 10.1007/s10142-016-0510-3

Merlot, S., Gosti, F., Guerrier, D., Vavasseur, A., and Giraudat, J. (2001). The ABI1 and ABI2 protein phosphatases 2C act in a negative feedback regulatory loop of the abscisic acid signalling pathway. *Plant J* 25**,** 295-303.doi: 10.1046/j.1365-313x.2001.00965.x

Mohanta, T.K., Park, Y.H., and Bae, H. (2016). Novel Genomic and Evolutionary Insight of WRKY Transcription Factors in Plant Lineage. *Sci Rep* 6**,** 37309.doi: 10.1038/srep37309

Moore, R.C., and Purugganan, M.D. (2005). The evolutionary dynamics of plant duplicate genes. *Curr Opin Plant Biol* 8**,** 122-128.doi: 10.1016/j.pbi.2004.12.001

Murat, F., Armero, A., Pont, C., Klopp, C., and Salse, J. (2017). Reconstructing the genome of the most recent common ancestor of flowering plants. *Nat Genet* 49**,** 490-496.doi: 10.1038/ng.3813

Narusaka, M., Shirasu, K., Noutoshi, Y., Kubo, Y., Shiraishi, T., Iwabuchi, M., et al. (2009). RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *Plant J* 60**,** 218-226.doi: 10.1111/j.1365-313X.2009.03949.x

Nehrt, N.L., Clark, W.T., Radivojac, P., and Hahn, M.W. (2011). Testing the ortholog conjecture with comparative functional genomic data from mammals. *PLoS Comput Biol* 7**,** e1002073.doi: 10.1371/journal.pcbi.1002073

Okonechnikov, K., Golosova, O., Fursov, M., and Team, U. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 28**,** 1166-1167.doi: 10.1093/bioinformatics/bts091

Panchy, N., Lehti-Shiu, M., and Shiu, S.H. (2016). Evolution of Gene Duplication in Plants. *Plant Physiol* 171**,** 2294-2316.doi: 10.1104/pp.16.00523

Pandey, S.P., Roccaro, M., Schon, M., Logemann, E., and Somssich, I.E. (2010). Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of Arabidopsis. *Plant J* 64**,** 912-923.doi: 10.1111/j.1365-313X.2010.04387.x

Pandey, S.P., and Somssich, I.E. (2009). The role of WRKY transcription factors in plant immunity. *Plant Physiol* 150**,** 1648-1655.doi: 10.1104/pp.109.138990

Phukan, U.J., Jeena, G.S., and Shukla, R.K. (2016). WRKY Transcription Factors: Molecular Regulation and Stress Responses in Plants. *Front Plant Sci* 7**,** 760.doi: 10.3389/fpls.2016.00760

Rinerson, C.I., Rabara, R.C., Tripathi, P., Shen, Q.J., and Rushton, P.J. (2015). The evolution of WRKY transcription factors. *BMC Plant Biol* 15**,** 66.doi: 10.1186/s12870-015-0456-y

Roulin, A., Auer, P.L., Libault, M., Schlueter, J., Farmer, A., May, G., et al. (2013). The fate of duplicated genes in a polyploid plant genome. *Plant J* 73**,** 143-153.doi: 10.1111/tpj.12026

Rushton, P.J., Somssich, I.E., Ringler, P., and Shen, Q.J. (2010). WRKY transcription factors. *Trends Plant Sci* 15**,** 247-258.doi: 10.1016/j.tplants.2010.02.006

Sahebi, M., Hanafi, M.M., Van Wijnen, A.J., Rice, D., Rafii, M.Y., Azizi, P., et al. (2018). Contribution of transposable elements in the plant's genome. *Gene* 665**,** 155-166.doi: 10.1016/j.gene.2018.04.050

Sanchez-Sevilla, J.F., Vallarino, J.G., Osorio, S., Bombarely, A., Pose, D., Merchante, C., et al. (2017). Gene expression atlas of fruit ripening and transcriptome assembly from RNA-seq data in octoploid strawberry (Fragaria x ananassa). *Sci Rep* 7**,** 13737.doi: 10.1038/s41598-017-14239-6

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463**,** 178-183.doi: 10.1038/nature08670

Shulaev, V., Sargent, D.J., Crowhurst, R.N., Mockler, T.C., Folkerts, O., Delcher, A.L., et al. (2011). The genome of woodland strawberry (Fragaria vesca). *Nat Genet* 43**,** 109-116.doi: 10.1038/ng.740

Singh, A., Sharma, A.K., Singh, N.K., Sonah, H., Deshmukh, R., and Sharma, T.R. (2019). Understanding the Effect of Structural Diversity in WRKY Transcription Factors on DNA Binding Efficiency through Molecular Dynamics Simulation. *Biology (Basel)* 8.doi: 10.3390/biology8040083

Solovyev, V., Kosarev, P., Seledsov, I., and Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol* 7 Suppl 1**,** S10 11-12.doi: 10.1186/gb-2006-7-s1-s10

Sperschneider, J., Catanzariti, A.M., Deboer, K., Petre, B., Gardiner, D.M., Singh, K.B., et al. (2017). LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Sci Rep* 7**,** 44598.doi: 10.1038/srep44598

Sun, Y., and Yu, D. (2015). Activated expression of AtWRKY53 negatively regulates drought tolerance by mediating stomatal movement. *Plant Cell Reports* 34**,** 1295-1306.doi: 10.1007/s00299-015-1787-8

Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* 34**,** W609-612.doi: 10.1093/nar/gkl315

Symons, G.M., Chua, Y.J., Ross, J.J., Quittenden, L.J., Davies, N.W., and Reid, J.B. (2012). Hormonal changes during non-climacteric ripening in strawberry. *J Exp Bot* 63**,** 4741-4750.doi: 10.1093/jxb/ers147

Tennessen, J.A., Govindarajulu, R., Ashman, T.L., and Liston, A. (2014). Evolutionary origins and dynamics of octoploid strawberry subgenomes revealed by dense targeted capture linkage maps. *Genome Biol Evol* 6**,** 3295-3313.doi: 10.1093/gbe/evu261

Tiley, G.P., Ane, C., and Burleigh, J.G. (2016). Evaluating and Characterizing Ancient Whole-Genome Duplications in Plants with Gene Count Data. *Genome Biol Evol* 8**,** 1023-1037.doi: 10.1093/gbe/evw058

Toljamo, A., Blande, D., Karenlampi, S., and Kokko, H. (2016). Reprogramming of Strawberry (Fragaria vesca) Root Transcriptome in Response to Phytophthora cactorum. *PLoS One* 11**,** e0161078.doi: 10.1371/journal.pone.0161078

Tsuda, K., and Somssich, I.E. (2015). Transcriptional networks in plant immunity. *New Phytol* 206**,** 932-947.doi: 10.1111/nph.13286

Van Verk, M.C., Bol, J.F., and Linthorst, H.J. (2011). WRKY transcription factors involved in activation of SA biosynthesis genes. *BMC Plant Biol* 11**,** 89.doi: 10.1186/1471-2229-11-89

Van Verk, M.C., Pappaioannou, D., Neeleman, L., Bol, J.F., and Linthorst, H.J. (2008). A Novel WRKY transcription factor is required for induction of PR-1a gene expression by salicylic acid and bacterial elicitors. *Plant Physiol* 146**,** 1983-1995.doi: 10.1104/pp.107.112789

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., et al. (2010). The genome of the domesticated apple (Malus x domestica Borkh.). *Nat Genet* 42**,** 833-839.doi: 10.1038/ng.654

Vicient, C.M., and Casacuberta, J.M. (2017). Impact of transposable elements on polyploid plant genomes. *Ann Bot* 120**,** 195-207.doi: 10.1093/aob/mcx078

Villanueva-Canas, J.L., Laurie, S., and Alba, M.M. (2013). Improving genome-wide scans of positive selection by using protein isoforms of similar length. *Genome Biol Evol* 5**,** 457-467.doi: 10.1093/gbe/evt017

Voorrips, R.E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 93**,** 77-78.doi: 10.1093/jhered/93.1.77

Wang, D., Amornsiripanitch, N., and Dong, X. (2006). A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants. *PLoS Pathog* 2**,** e123.doi: 10.1371/journal.ppat.0020123

Wang, F., Zhang, F., Chen, M., Liu, Z., Zhang, Z., Fu, J., et al. (2017). Comparative Transcriptomics Reveals Differential Gene Expression Related to Colletotrichum gloeosporioides Resistance in the Octoploid Strawberry. *Front Plant Sci* 8**,** 779.doi: 10.3389/fpls.2017.00779

Wang, Q., Wang, M., Zhang, X., Hao, B., Kaushik, S.K., and Pan, Y. (2011). WRKY gene family evolution in Arabidopsis thaliana. *Genetica* 139**,** 973-983.doi: 10.1007/s10709-011-9599-4

Wei, W., Cui, M.Y., Hu, Y., Gao, K., Xie, Y.G., Jiang, Y., et al. (2018). Ectopic expression of FvWRKY42, a WRKY transcription factor from the diploid woodland strawberry (Fragaria vesca), enhances resistance to powdery mildew, improves osmotic stress resistance, and increases abscisic acid sensitivity in Arabidopsis. *Plant Sci* 275**,** 60-74.doi: 10.1016/j.plantsci.2018.07.010

Wei, W., Hu, Y., Han, Y.T., Zhang, K., Zhao, F.L., and Feng, J.Y. (2016). The WRKY transcription factors in the diploid woodland strawberry Fragaria vesca: Identification and expression analysis under biotic and abiotic stresses. *Plant Physiol Biochem* 105**,** 129-144.doi: 10.1016/j.plaphy.2016.04.014

Xiang, Y., Huang, C.H., Hu, Y., Wen, J., Li, S., Yi, T., et al. (2017). Evolution of Rosaceae Fruit Types Based on Nuclear Phylogeny in the Context of Geological Times and Genome Duplication. *Mol Biol Evol* 34**,** 262-281.doi: 10.1093/molbev/msw242

Yang, Y., Zhou, Y., Chi, Y., Fan, B., and Chen, Z. (2017). Characterization of Soybean WRKY Gene Family and Identification of Soybean WRKY Genes that Promote Resistance to Soybean Cyst Nematode. *Sci Rep* 7**,** 17804.doi: 10.1038/s41598-017-18235-8

Yang, Z., and Nielsen, R. (2002). Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol* 19**,** 908-917.doi: 10.1093/oxfordjournals.molbev.a004148

Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., et al. (2018). WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. *Nucleic Acids Res* 46**,** W71-W75.doi: 10.1093/nar/gky400

Zhang, C.Q., Xu, Y., Lu, Y., Yu, H.X., Gu, M.H., and Liu, Q.Q. (2011). The WRKY transcription factor OsWRKY78 regulates stem elongation and seed development in rice. *Planta* 234**,** 541-554.doi: 10.1007/s00425-011-1423-y

Zhang, L., Chen, L., and Yu, D. (2018a). Transcription Factor WRKY75 Interacts with DELLA Proteins to Affect Flowering. *Plant Physiol* 176**,** 790-803.doi: 10.1104/pp.17.00657

Zhang, L., Huang, X., He, C., Zhang, Q.Y., Zou, X., Duan, K., et al. (2018b). Novel Fungal Pathogenicity and Leaf Defense Strategies Are Revealed by Simultaneous Transcriptome Analysis of Colletotrichum fructicola and Strawberry Infected by This Fungus. *Front Plant Sci* 9**,** 434.doi: 10.3389/fpls.2018.00434

Zhang, Y., and Wang, L. (2005). The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evol Biol* 5**,** 1.doi: 10.1186/1471-2148-5-1

Zheng, Z., Mosher, S.L., Fan, B., Klessig, D.F., and Chen, Z. (2007). Functional analysis of Arabidopsis WRKY25 transcription factor in plant defense against Pseudomonas syringae. *BMC Plant Biol* 7**,** 2.doi: 10.1186/1471-2229-7-2

Zheng, Z., Qamar, S.A., Chen, Z., and Mengiste, T. (2006). Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant J* 48**,** 592-605.doi: 10.1111/j.1365-313X.2006.02901.x

Zhou, H., Li, Y., Zhang, Q., Ren, S., Shen, Y., Qin, L., et al. (2016). Genome-Wide Analysis of the Expression of WRKY Family Genes in Different Developmental Stages of Wild Strawberry (Fragaria vesca) Fruit. *PLoS One* 11**,** e0154312.doi: 10.1371/journal.pone.0154312

Capítulo 3:

The vq motif-containing proteins in the diploid and octoploid strawberry.

Parte de este capítulo ha sido publicado el 20 de marzo de 2019:

Garrido-Gala J, Higuera JJ, Muñoz-Blanco J, Amil-Ruiz F, Caballero JL. *The VQ motif-containing proteins in the diploid and octoploid strawberry*. Sci Rep. 2019 Mar 20;9(1):4942. doi: 10.1038/s41598-019-41210-4.

Capítulo 3: The VQ motif-containing proteins in the diploid and octoploid strawberry.

3.1. Abstract

The plant VQ motif-containing proteins are a recently discovered class of plant regulatory proteins interacting with WRKY transcription factors capable of modulate their activity as transcriptional regulators. The short VQ motif (FxxhVQxhTG) is the main element in the WRKY-VQ interaction, whereas a newly identified variable upstream amino acid motif appears to be determinant for the WRKY specificity. The VQ family has been studied in several species and seems to play important roles in a variety of biological processes, including response to biotic and abiotic stresses. Here, we present a systematic study of the VQ family in both diploid (*Fragaria vesca*) and octoploid (*Fragaria x ananassa*) strawberry species. Thus, twenty-five VQ-encoding genes were identified and twenty-three were further confirmed by gene expression analysis in different tissues and fruit ripening stages. Their expression profiles were also studied in *F. ananassa* fruits affected by anthracnose, caused by the ascomycete fungus *Colletotrichum*, a major pathogen of strawberry, and in response to the phytohormones salicylic acid and methyl-jasmonate, which are well established as central stress signals to regulate defence responses to pathogens. This comprehensive analysis sheds light for a better understanding of putative implications of members of the VQ family in the defence mechanisms against this major pathogen in strawberry.

3.2. Introduction

Plant growth and development are constantly affected by changing environmental conditions and stresses in both, natural and agricultural settings (Garner et al., 2016). Among the most important biotic stresses are those caused by microbial pathogens like viruses, bacteria and fungi. Thus, plants have evolved various and complex defence systems to protect themselves from pathogens, which are finely regulated by a large and diverse set of regulatory proteins including transcription factors (TFs), that bind to specific *cis*-regulatory elements in the promoter region of target genes controlling their transcription (Alves et al., 2014).

Several TF families are particularly involved in regulating the defence responses in plants: AP2/ARF, bHLH, bZip, MYB, NAC and WRKY (Tsuda and Somssich, 2015). WRKY TFs are one of the largest families of transcriptional regulators existing in plants. They are structurally characterized by a highly conserved DNA binding domain, about 60 amino acids long, harbouring one or two core motif WRKYGQK and a Zinc-finger-like motif with two variants: C2H2 or C2HC. WRKY proteins are classified in Groups I, II and III on the basis of both the number of WRKYGQK motifs and the features of their Zinc-finger motif. The polyphyletic Group II is further divided into subgroups IIa, IIb, IIc, IId

and IIe based on differences in their amino acid sequence (Eulgem et al., 2000;Rushton et al., 2010;Jiang et al., 2017). WRKY TFs bind to target gene *cis*-elements with sequence TTGAC[C/T] known as W-box. They have been referred as the "jack of many trades in plants", forming an intricate network that play diverse roles in regulating the transcriptional activity of plant cells to execute several developmental programmes and stress responses(Bakshi and Oelmuller, 2014).

Regulatory proteins that do not bind DNA directly can form protein-protein complexes with TFs to fine-tune the transcriptional response. A recently identified group of proteins containing a short and conserved amino acid motif, named VQ proteins, belongs to this class. VQ motif-containing proteins have been characterized in a number of plants including Arabidopsis with 34 VQ members (Cheng et al., 2012), rice with 40 VQ members (Li et al., 2014a), soybean with 74 VQ members (Zhou et al., 2016), and grapevine with 18 VQ members (Wang et al., 2015b). Very recently, an *in silico* analysis of the VQ protein family addressing the phylogenetic relationships and microevolution of VQ genes in the genus *Fragaria* was published (Zhong et al., 2018). Low sequence similarity has been found between known plant VQ proteins and proteins from other organism, suggesting that the VQ family is highly specific to plants(Jing and Lin, 2015). However, in a more recent phylogenetic study, some non-plant proteins containing partial or divergent VQ motifs were found (Jiang et al., 2018). Plant VQs are relatively short in length and mostly coded by intronless genes but two or more introns have also been found (Li et al., 2014b). All plant VQs share the conserved amino acid motif FxxhVQxhTG involved in protein binding with several WRKY TFs but they are variable in length and highly divergent in amino acid sequence outside of the VQ domain (Cheng et al., 2012).

A number of evidences have shown that plant VQ proteins bind to the C-terminal WRKY domain of Group I and to the single one of Group IIc WRKY proteins. Lai et al. (2011) demonstrated that the conserved V and Q residues in the VQ motif of SIB1 (VQ23) are essential for the interaction with WRKY33 in Arabidopsis and proposed the VQ motif as the core of the WRKY-interacting motif (Lai et al., 2011). Recently, Zhou et al. (2016) identified an additional eight amino acids long upstream motif close to the VQ motif as an important region for the affinity and specificity of WRKY-VQ binding. Thus, single amino acid substitutions changed the specificity patterns for their WRKY target in the mutated VQs while did not abolish their binding to WRKY proteins, except by deletion of the whole motif in GmVQ22 (Zhou et al., 2016). A very recent work in apple confirms that the amino acid residues flanking the core VQ motif are also required for the interaction with the WRKY domain, as well as the specificity of the VQ proteins for the Group I and IIc of WRKY TF (Dong et al., 2018).

VQ proteins can activate or repress transcriptional activity of Group I and IIc WRKY proteins(Jing and Lin, 2015). *Arabidopsis* VQ23 and VQ16 (SIB1 and SIB2, respectively) recognize the C-terminal WRKY domain of WRKY33 and stimulate its DNA binding activity, thus playing a positive role in defence against necrotrophs. Indeed, WRKY33, SIB1 and SIB2 were markedly induced by *B. cinerea* infection and showed similar

expression patterns (Lai et al., 2011). Arabidopsis VQ29 can interact with WRKY25 and WRKY33 (Cheng et al., 2012) but also with PIF1, a bHLH TF that negatively regulates light-dependent seed germination (Li et al., 2014b). While it positively stimulates the activity of PIF1 (hence, acting as repressor of seedling de-etiolation), its overexpression increases susceptibility to *B. cinerea* in *Arabidopsis* leaves and thus, acting as a negative regulator of the defence response (Wang et al., 2015a). In contrast, *AtVQ29* expression is induced early upon infection in Arabidopsis roots by *P. parasitica* and it is required to restrict the pathogen development independently of SA-, JA- and ET-mediated defence activation(Le Berre et al., 2017). Thus, AtVQ29 can modulate different responses either positively or negatively, depending of the tissue and stimulus. On the other hand, MKS1 (AtVQ21) forms complexes with MPK4 and WRKY33. After pathogen infection, phosphorylation by MPK4 occurs and the MKS1-WRKY33 complex is activated and can dissociate, promoting the *PAD3* expression and camalexin biosynthesis. Thus, MKS1 was shown to be acting as repressor in absence of *P. syringae* infection or flagelin treatment (Qiu et al., 2008). Similarly to the MKS1-MPK4 interaction, other VQ proteins have been proven as substrates for the mitogen-activated protein kinases MPK3 and MPK6. Phosphorylation of a subset of VQ proteins by MPK3/6 destabilizes their association with WRKY proteins, leading to their release from the complex. This can promote direct WRKY-DNA binding or alternatively, the substitution for a different VQ protein, thus regulating transcriptional activity upon PAMP recognition and providing an additional and finely regulated mechanism for transcriptional plant defence control (Pecher et al., 2014;Weyhe et al., 2014).

The genus *Fragaria* (*Rosaceae*) comprises about 24 species worldwide distributed, including diploids and polyploids, wild and cultivated members(DiMeglio et al., 2014). The octoploid cultivated strawberry (*Fragaria x ananassa* Duch.) is an important and appreciated fruit crop with valuable organoleptic and nutritional qualities (Schwab et al., 2009). With an annually increasing, world production of more than 9 Mt and 401,862 ha cultivated (FAOSTAT, 2016) it is presumably, the most economically important soft berry.

For the past several years, continuous efforts to dissect the genomic structure and genetic regulation of this valuable crop are being made, which is an inescapable requisite for both, basic and applied research, in aspects such as identification of genetic markers linked to valuable traits for breeding, improving fruit quality, stress response regulation or comparative genomics. The diploid *Fragaria vesca* genome was firstly sequenced and published in 2011 (Shulaev et al., 2011) and recently re-sequenced *de novo* (Edger et al., 2018). The first assembly versions have been later improved (Tennessen et al., 2014) and re-annotated subsequently with RNA-seq support (Darwish et al., 2015;Li et al., 2018). The genomes of cultivated strawberry (*Fragaria* x ananassa) and four wild *Fragaria* species, representing the genetic diversity into the genus (*F. nipponica, F. iinumae, F. orientalis,* and *F. nubicola*) were sequenced and a reference genome of *Fragaria* x *ananassa* was partially assembled and designated as FANhybrid_r1.2 (Hirakawa et al., 2014). One remarkable result was the high sequence similarity with *F. vesca* (assembly version 1.1), with only 5.23% of non-

homologous sequences. Besides the previously observed high levels of conserved macrosynteny and colinearity between the diploid and octoploid *Fragaria* genomes, this find strengthen support for the use of the former as model for genomic research in the latter, much more complex (Rousseau-Gueutin et al., 2009).

Strawberry anthracnose is a severe disease caused by the fungus *Colletotrichum*, one of the most important genera of strawberry pathogens. *Colletotrichum acutatum*, *Colletotrichum fragariae* and *Colletotrichum gloeosporioides* are the three major species causing the fruit and crown rot diseases in strawberry (Denoyes-Rothan et al., 2003). *Colletotrichum spp*. is a hemibiotrophic pathogen, switching from a short, symptomless biotrophic stage to a necrotrophic phase, characterized by host cell death and extensive tissue colonization, producing the typical anthracnose symptoms. In the past years, a few transcriptomic studies focused in the *Colletotrichum*-strawberry interaction have been conducted (Guidarelli et al., 2011;Amil-Ruiz et al., 2016;Wang et al., 2017a) with no specific findings about possible roles of the VQ protein family in the defence mechanisms of strawberry to this pathogen.

In the present work, a deep and systematic analysis of the strawberry VQ proteins was performed using the latest available, annotated reference genomes and transcriptomes of both, diploid and octoploid species. We investigated the phylogenetic relationships, conserved domains and structure among four representative species, in an effort to establish a more robust classification of the VQ family into functional groups. As a remarkable result of the domains screening, we also describe a novel protein domain association and propose the name of R protein-VQ for the VQ proteins harbouring NBS-ARC and LRR-like domains, typical of R proteins. In addition, in order to elucidate putative roles of members of the VQ protein family in the defence mechanisms of strawberry against pathogens, the expression pattern of the VQ genes (*Fragaria* x ananassa cv. Camarosa) were analysed in different plant tissues and fruit ripening stages as well as in response to fruit infection with *C. acutatum* and salicylic acid and methyl-jasmonate treatments.

3.3. Materials and Methods

3.3.1. Plant material and treatments

Field-grown strawberry plants (*Fragaria* x ananassa Duch cv. Camarosa) tissues and fruits were collected for gene expression analysis: roots, crown, petiole, leaf, mature flowers and fruits (receptacles and achenes) in different developmental stages. Naturally infected red strawberry fruits (cv. Camarosa) showing symptoms of anthracnose were collected and pooled into four groups representing increasing stages of disease development as described (Encinas-Villarejo et al., 2009): control (healthy) and grades 1 to 3 of infection (G1 to G3, corresponding to ID1 to ID3 in the referred paper). For phytohormone treatments, four weeks old *in vitro* Camarosa plants,

growing in solid N30K medium (Margara, 1989), were sprayed with mock, 5 mM salicylic acid (SA) or 2 mM methyl jasmonate (MeJA). Replicates were harvested at 6, 12 and 24 hours. All samples were frozen in liquid nitrogen and stored at -80 ºC until RNA purification.

3.3.2. Identification and molecular characterization of the VQ Family in *F. vesca* and *F. ananassa*

The previously published *Arabidopsis thaliana*, soybean and grapevine sequences of VQ proteins used in this study were downloaded from their original sources (Cheng et al., 2012;Wang et al., 2015b;Zhou et al., 2016). *Fragaria vesca* (genome v2.0.a2 and v4.0.a1) and *Fragaria* x ananassa (FANhybrid_r1.2 and GDR *Fragaria x ananassa* RefTrans V1) sequences are available online in the Genome Database for Rosaceae website (https://www.rosaceae.org/). The Hidden Markov Model (HMM) of the VQ motif family (PF05678) was downloaded from the Pfam 31.0 database (https://pfam.xfam.org/) and used as query in HMMER3 search, performed in the freeware tool UGENE v1.21(Okonechnikov et al., 2012). The candidate sequences were confirmed to include the VQ conserved motif using Pfam 31.0 and the NCBI's Conserved Domain Database (CDD) (Marchler-Bauer et al., 2017). Additional domains found were also checked using the Conserved Domain Architecture Retrieval Tool (CDART). Selected features of the *F. vesca VQ* (*FvVQ*) genes, protein length, ORF length, molecular weight (kDa) and isoelectric point (PI) of each gene were calculated using ExPASy (https://web.expasy.org/). Nuclear location signals (NLSs) were detected by SeqNLS (Lin and Hu, 2013) and subcellular locations were predicted by LOCALIZER (Sperschneider et al., 2017). The chromosomal distribution of the *FvVQ* genes was drawn with Mapchart 2.32 (Voorrips, 2002) and the intron-exon structure by using the online tool GSDS 2.0 (Hu et al., 2015). The 1500 bp tracks upstream of *FvVQ*s (or less, if overlapped a previous coding sequence) were screened for known regulatory *cis*-elements and transcription factor binding sites (TFBS) at PLANTCARE and PLANTPAN 2.0 websites. Finally, *the novo* motif discovery was carried out by the MEME Suite v5.0.1 (http://meme-suite.org/tools/meme) on the FvVQ proteins with the following modified parameters: maximum number of motifs set to 20; motif width: 6 to 50.

3.3.3. Multiple sequence alignment and phylogenetic analysis

Structure-based alignment of *Arabidopsis*, soybean, grapevine and *F. vesca* VQ domain sequences was performed in PROMALS3D to gain additional information about the protein secondary structure (Pei et al., 2008). This alignment was loaded in MEGA 7.0 (Kumar et al., 2016) as basis to construct unrooted phylogenetic trees by the neighbor-joining (N-J) method with 1000 bootstrap replicates. Evolutionary distances were computed by the p-distance method and pairwise deletion. Orthologous sequences between strawberry and *Arabidopsis* were identified by their best hit in reciprocal BLASTP searches and checked in OrthoMCL DB (http://orthomcl.org/orthomcl/). The

homology with *Arabidopsis* proteins, was applied in the construction of a protein functional association network by STRING 10.5 (https://string-db.org/) with medium confidence (0.400).

3.3.4. Quantitative RT-PCR expression analysis

Specific primer pairs (Table S1) were designed by Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Optimal annealing temperatures were assessed by gradient-PCR. Total RNA extraction from strawberry samples, further purification, quality checks, cDNA synthesis, PCR efficiency determination and RT-qPCR runs were carried out as previously described (Amil-Ruiz et al., 2013). Reference genes used in this work were *FaEF1α* and *FaACTIN* for tissues and infected fruits, or *FaEF1α* and *FaGAPDH2* for *in vitro* cultured plants (Amil-Ruiz et al., 2013). Normalized relative quantification of the gene expression was calculated using the Hellemans' modification of the Pfaffl method to include several reference genes (Pfaffl, 2001;Hellemans et al., 2007). Gene expression in different tissues was calculated as $2^{-\Delta Ct}$ (ΔCt= mean Ct of *VQ* gene minus the geometric mean of reference genes) and represented as heatmap, generated by Heatmapper (Babicki et al., 2016) with complete linkage and Spearman Rank Correlation settings. A Venn diagram of the up-regulated *FaVQ* for every treatment was generated by BioVenn (Hulsen et al., 2008).

3.3.5. Statistical analysis

Data were analysed in Microsoft Excel, using the Real Statistics Resource Pack software, release 5.4 (http://www.real-statistics.com/). All data were tested for normality using Shapiro-Wilk test (α=0.05). One-way ANOVA, followed by Dunnett's test (two-tailed) post-hoc were performed at α=0.05 and 0.01. Three biological replicates were used: three fruits sharing the same symptoms or the same developmental stage, or three *in vitro* plantlets form a biological replicate (n=3). Mean, standard error (SE) and significant differences of *P≤0.05 and **P≤0.01 are represented in figures.

Table S1. Primers used in this study

Gene	Forward primer	Reverse primer	Annealing (ºC)
VQ1	TCGTGACCTTGTCCAGAAGC	CCGAAACCGGAGAAGGAACA	62
VQ2	TGATTGATTCGAGGGAGGCG	GCAGTCACCTCCCATTTCCA	58
VQ3	AAATCAAGAAGCCACCGCCT	GCGTGGATGACTTTGGGAGA	62
VQ4	TCGGAACACCAACCCGAAAA	ACTCGTAAATGGGGGTGCAG	62
VQ5	GGTGGGGAATCGACAATGGA	TGAAGGAATCGGAGGAAGCG	58
VQ6	GGAGATTTTGTCCCCGAGCA	ACGGGTGCAAGTAAAACCCT	62
VQ7	TTCCCCAGTAGTACCAGCCA	TTGTTGGCGTTGGCGTTAAG	62
VQ8	GGTGCAGAAGATGACCGGAA	TACTGCGACGCATTGATCCT	58
VQ9	GTGACCCCATTGAACGACGA	CCGGAAACAATGGCAGAAGC	62
VQ10	GCAGCTTACTGGTTCCCCAT	GGGAGGCCTAACAAAGCCAT	62
VQ11	TAAACAGCAGCCTTCTCGCA	TGTGTGCGGCTGATTCGTTA	62
VQ12	GCCACCCGGTCATCATCTAC	GCCTACGTTGTTGTTGGTGC	58
VQ13	CACCGGAAAGAGCCTTCCTT	AACTCGTCCGGTAAACCGTC	62
VQ14	CGATTCATCCGCCGAAAACC	GCCCATAGGCGGAATGAGAG	62
VQ15	CTCCCACGACTGTTCTGACC	CCCGAAGATACTGTCGAGCC	62
VQ16	AGCCCTATGATGGTTCAAGCC	TTCATCAGCTCATCGCCGTT	62
VQ17	GACGGGAGTGAGGTTCAGTG	GGTCAAACAGAAACGCCGAC	62
VQ18	CGCCAGATCTCCACCCTACT	GGCACAAAGCTGTCAAGCAT	62
VQ19	ATACACAGTTGCAGGGTCCG	TCATGATTCCCGGTAAGGCG	62
VQ20	AGACTCACCGGAAAGCCATC	CCGCCTCCGTTACTTCTCAG	62
VQ21	CCTTCAAGCTCCACGAGAGG	CGTGCCAGAAACTCAAACGG	62
VQ22	GCTACCTCCTCATCACACCC	CCACAGCCCGGAAGAATCA	62
VQ23	CAAGACCCTTGTGCAAGACC	AGTACTCCTTTCAGGCTCCAC	62
VQ24	TACGGCGTCCTTTACAACCC	ATGATTGTACCCAGGTGGCG	62
VQ25	AGGTTCAGACTAGCGCCTCT	TCGGCTTCCTTGAGTTGCTT	62
FaACTIN	GGGCCAGAAAGATGCTTATGTCGG	GGGCAACACGAAGCTCATTGTAGAAG	65
FaEF1α	TGGATTTGAGGGTGACAACATGA	GTATACATCCTGAAGTGGTAGACGGAGG	65
FaGAPDH2	GGCTTCTATCTCAACCGGCTCGTCTT	CTTCCCACTGCTCCCTGATCTCTGATAC	65

3.4. Results and Discussion

3.4.1. VQ members of *F. vesca* and *F. ananassa*

Characteristics of FvVQ genes found in the genome annotation of *F. vesca* v2.0.a2 and their corresponding predicted proteins are listed in Table 1 and named *FvVQ1* to *FvVQ25* according to their chromosomal order. Most of the *FvVQ* genes are relatively short, intronless (18 out of 25) and they are unevenly distributed in the genome, being the chromosome 6, which contains the higher number of them (Figure 1). As it has been previously described for other species (Cheng et al., 2012;Wang et al., 2015b;Zhou et al., 2016), most of the strawberry *FvVQ* genes encode short proteins ranging from 131 to 678 aa, with an average of 274. Also, two and five alternative RNA splicing forms were found for *genes FvVQ6* and *FvVQ13*, respectively. Furthermore, as expected, most of the FvVQ proteins are predicted to locate in the nucleus and show monopartite or bipartite nuclear location signals, while FvVQ1 and FvVQ3 also contain putative chloroplast import signals. Despite of the absence of classical NLSs in some FvVQ, their predictions agree with previous evidence that VQ proteins can be located in the nucleus, where they interact with WRKY TFs to modulate the gene expression (Kim et al., 2013).

BLASTN searches with the *FvVQ* set as query were performed to find the *Fragaria* x ananassa orthologous *VQ* genes (*FaVQ*) and transcripts from both the reference genome and RefTrans V1 RNA-seq annotation. Results were compared with those obtained by the HMM3 searches for both datasets. Thus, many orthologous sequences were found matching with the RefTrans V1 transcripts and the FANhybrid scaffolds and their predicted genes, with the exceptions of *FvVQ2*, *FvVQ22* and *FvVQ23* (Table 2). Interestingly, PCR amplifications of genomic DNA and cDNA with the specific designed primers showed that strawberry *VQ22* and *VQ23* failed to be detected in cDNA samples from both the diploid *F. vesca* cv. Reina de los Valles and *F.* ananassa cv. Camarosa in all tested tissues, developmental stages and treatment conditions (see further below). Using BLASTN, we also compared the curated *VQ* genes found by HMM3 on the Fv v2.0.a2 as well as the newest Fv v4.0.a1 genome annotation, with negative results for these two genes (Table S2). Thus, we think that prediction of these two genes should be taken with caution.

Name	Gene id. (*F. vesca* v2.0.a2)	Genomic location	Splicing forms	Exons	Protein properties				
					Length (aa)	MW (KDa)	PI	Domains	Subcellular location
FvVQ1	*gene05795-v1.0-hybrid*	Fvb1:11109731..11110700 (+ strand)		1	177	19,4	9,58	VQ	ch *
FvVQ2	*gene12350-v1.0-hybrid*	Fvb1:12947350..12948422 (+ strand)		2	186	20,69	9,55	VQ	
FvVQ3	*gene17872-v1.0-hybrid*	Fvb1:14815114..14815929 (- strand)		1	271	29,32	8,10	VQ	ch/n
FvVQ4	*gene16338-v1.0-hybrid*	Fvb1:17333502..17343519 (- strand)		2	460	49,20	6,67	VQ	n *
FvVQ5	*gene32666-v1.0-hybrid*	Fvb2:9213921..9216468 (+ strand)		3	361	39,48	6,87	VQ	n *
FvVQ6	*gene02378-v2.0.a2-hybrid*	Fvb3:5205231..5207960 (- strand)	t1, t2	3	254	27,71	9,76	VQ	*
FvVQ7	*gene19985-v1.0-hybrid*	Fvb3:7601133..7601903 (+ strand)		1	238	26,20	5,75	VQ	*
FvVQ8	*gene02352-v2.0.a2-hybrid*	Fvb3:8749508..8750454 (+ strand)		2	148	16,35	10,14	VQ	
FvVQ9	*gene17096-v2.0.a2-hybrid*	Fvb4:8384711..8386867 (+ strand)		1	261	28,78	9,58	VQ	
FvVQ10	*gene22762-v1.0-hybrid*	Fvb4:23658011..23658958 (+ strand)		1	315	33,98	10,80	VQ	
FvVQ11	*gene03724-v1.0-hybrid*	Fvb4:27000140..27000874 (- strand)		1	244	27,19	6,82	VQ	
FvVQ12	*gene25153-v1.0-hybrid*	Fvb5:6972449..6973138 (- strand)		1	229	24,75	5,83	VQ	n *
FvVQ13	*gene26510-v2.0.a2-hybrid*	Fvb5:25430959..25435652 (+ strand)	t5	2	676	77,38	7,53	VQ, LRR8, NB-ARC	
			t2	3	676	77,39	7,54		
			t4	3	678	77,27	6,89		
			t1,t3	4	678	77,28	6,90		
FvVQ14	*gene11973-v1.0-hybrid*	Fvb5:26904474..26905877 (+ strand)		1	330	35,67	10,05	VQ	*
FvVQ15	*gene22130-v1.0-hybrid*	Fvb5:27727944..27729224 (- strand)		1	426	46,02	7,02	VQ	n *
FvVQ16	*gene16601-v1.0-hybrid*	Fvb6:1405916..1406311 (+ strand)		1	131	15,15	5,56	VQ	n *
FvVQ17	*gene00230-v1.0-hybrid*	Fvb6:4590905..4592326 (+ strand)		1	248	26,40	9,47	VQ	n *
FvVQ18	*gene21583-v1.0-hybrid*	Fvb6:12584223..12585325 (+ strand)		1	149	16,35	6,05	VQ	n *
FvVQ19	*gene18133-v1.0-hybrid*	Fvb6:17070375..17071259 (+ strand)		1	148	16,56	9,25	VQ	n *
FvVQ20	*gene22640-v1.0-hybrid*	Fvb6:18023651..18024721 (- strand)		1	184	20,68	9,27	VQ	n *
FvVQ21	*gene15994-v1.0-hybrid*	Fvb6:22449818..22451424 (+ strand)		1	195	21,36	8,03	VQ	
FvVQ22	*gene28736-v1.0-hybrid*	Fvb6:30523488..30524048 (+ strand)		1	186	20,78	4,30	VQ	n
FvVQ23	*gene28758-v1.0-hybrid*	Fvb6:30654008..30657346 (- strand)		3	263	29,62	4,77	VQ	n *
FvVQ24	*gene24219-v1.0-hybrid*	Fvb6:31699569..31700294 (+ strand)		1	241	25,53	10,24	VQ	n *
FvVQ25	*gene14221-v1.0-hybrid*	Fvb7:2184725..2185715 (+ strand)		1	132	15,20	7,31	VQ	n *

Table 1. List of the strawberry VQ members and selected properties. Subcellular locations predicted by LOCALIZER are: ch, chloroplast; n, nucleus. Nuclear location signals (NLS) predicted by SeqNLS with scores greater than 0.5 are marked by asterisks.

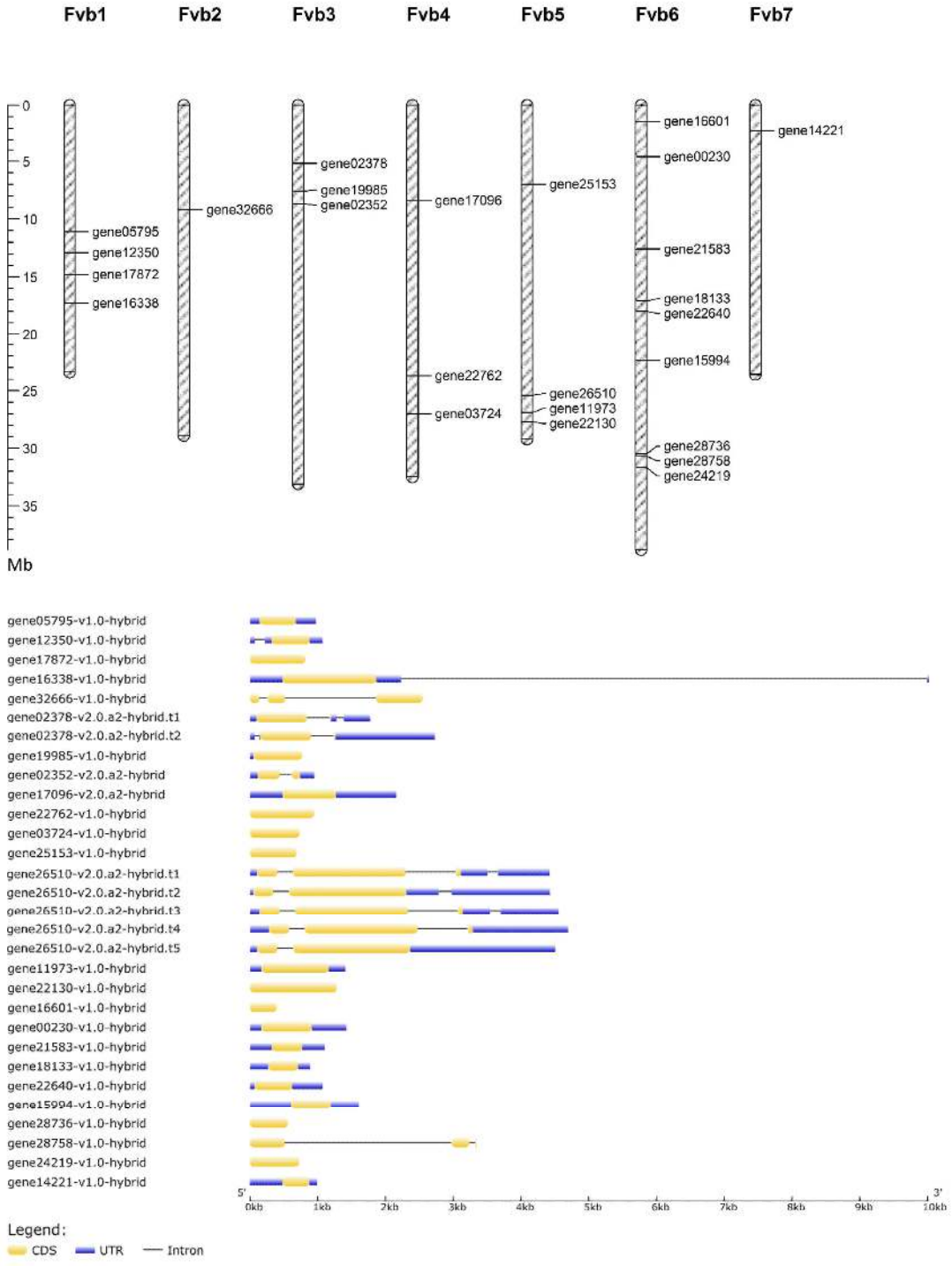


Figure 1. Chromosome mapping and gene structure of the Fragaria vesca VQ gene family.

query id	subject id	% identity	align length	mismatch	gap opens	q. start	q. end	s. start	s. end	e-value	bit score
FvH4_6g07790.1	gene00230-v1.0-hybrid	100	248	0	0	1	248	1	248	0,0	500
FvH4_3g16620.1	gene02352-v2.0.a2-hybrid	100	109	0	0	1	109	1	109	5,58e-79	221
FvH4_3g09030.1	gene02378-v2.0.a2-hybrid	99.6	255	0	1	1	255	1	254	0,0	502
FvH4_4g27220.1	gene03724-v1.0-hybrid	100	244	0	0	1	244	1	244	0,0	507
FvH4_1g18020.1	gene05795-v1.0-hybrid	100	177	0	0	1	177	1	177	2,18e-130	354
FvH4_5g36700.1	gene11973-v1.0-hybrid	100	330	0	0	1	330	1	330	0,0	647
FvH4_1g25800.1	gene12350-v1.0-hybrid	100	186	0	0	1	186	1	186	8,64e-141	381
FvH4_7g03520.1	gene14221-v1.0-hybrid	100	132	0	0	1	132	1	132	2,04e-99	272
FvH4_6g29370.1	gene15994-v1.0-hybrid	100	195	0	0	1	195	1	195	1,13e-144	392
FvH4_1g21110.1	gene16338-v1.0-hybrid	100	460	0	0	1	460	1	460	0,0	919
FvH4_6g02470.1	gene16601-v1.0-hybrid	100	131	0	0	1	131	1	131	7,00e-102	278
FvH4_3g19800.1	gene17096-v2.0.a2-hybrid	100	261	0	0	1	261	1	261	0,0	516
FvH4_1g24040.1	gene17872-v1.0-hybrid	100	229	0	0	1	229	43	271	1,78e-166	451
FvH4_6g21030.1	gene18133-v1.0-hybrid	100	148	0	0	1	148	1	148	1,49e-113	309
FvH4_3g12720.1	gene19985-v1.0-hybrid	100	238	0	0	1	238	1	238	0,0	487
FvH4_6g25410.1	gene21583-v1.0-hybrid	100	130	0	0	1	130	20	149	1,07e-98	271
FvH4_5g37690.1	gene22130-v1.0-hybrid	100	426	0	0	1	426	1	426	0,0	854
FvH4_6g22390.1	gene22640-v1.0-hybrid	100	184	0	0	1	184	1	184	3,23e-140	380
FvH4_4g21960.1	gene22762-v1.0-hybrid	100	299	0	0	1	299	17	315	0,0	561
FvH4_6g39500.1	gene24219-v1.0-hybrid	100	241	0	0	7	247	1	241	4,49e-177	478
FvH4_5g11980.1	gene25153-v1.0-hybrid	100	229	0	0	1	229	1	229	5,03e-169	456
FvH4_5g34700.1	gene26510-v2.0.a2-hybrid	100	678	0	0	1	678	1	678	0,0	1387
FvH4_2g14340.1	gene32666-v1.0-hybrid	100	218	0	0	1	218	144	361	5,45e-161	441
no hit	gene28736-v1.0-hybrid										
no hit	gene28758-v1.0-hybrid										

Table S2. Concordances by BLASTN between VQ genes found in the F. vesca genomes v2.0.a2 and v4.0.a1

On the other hand, several homologous transcripts were found matching with the RefTrans V1 transcripts for many *FaVQs* indicating the existence of more alternative splicing forms in the strawberry octoploid than in the diploid species (Table 2), Thus, two different transcripts were found for genes *FaVQ1, FaVQ3, FaVQ6, FaVQ9, FaVQ14 and FaVQ17*, and three different transcripts for gene *FaVQ13*. It is worth to note that two out of the three different *FaVQ13* transcripts code for proteins preserving the full predicted domains structure while the third one lacks the LRR8 domain (see below). These findings evidence differences in the control of mRNA maturation of these genes between the two strawberry species at post-transcriptional level that remains to be further studied.

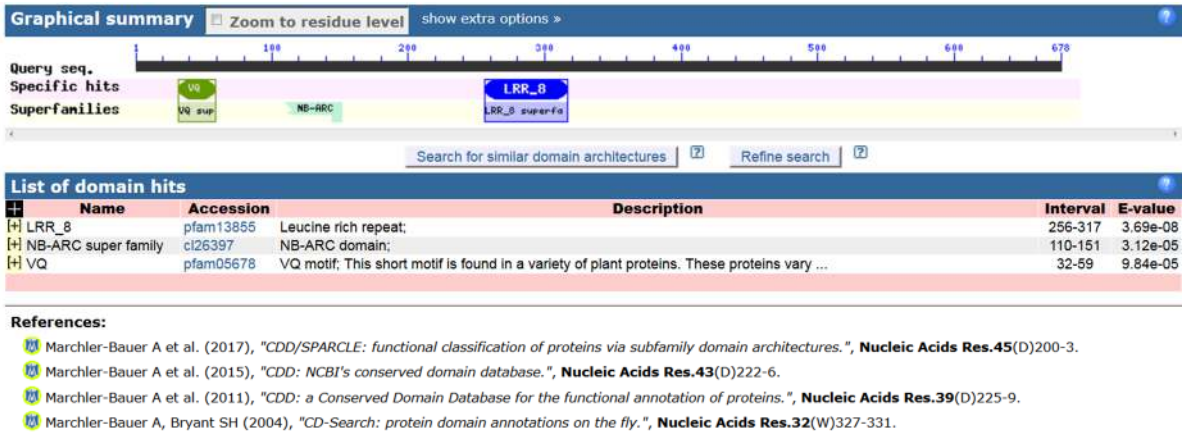
Curiously, Fv/FaVQ13 exhibited two domains found in the NBS-LRR class of R proteins: a NB-ARC domain (PF00931) and a Leu-rich motif (LRR_8; PF13855). This unusual association of such domains with the VQ motif is novel but a similar architecture is also found in CDD/CDART in several *Oryza sativa* proteins corresponding to OsVQ34 (Kim et al., 2013), *Beta vulgaris* (XP_010673130.1) and *Chenopodium quinoa* (XP_021745766.1) (Figure S1). Moreover, in a preliminary screening of the Phytozome 12.1 database we have found two more proteins, from *M. domestica* (MDP0000284090) and *C. grandiflora* (Cagra.22718s0002.1) harbouring similar domains. Interestingly, R protein domains have also been reported in combination with WRKY domains within the R protein-WRKY class of WRKY TF (Rushton et al., 2010;Rinerson et al., 2015). By analogy, we propose the name of R protein-VQ for this novel class of VQ proteins. Besides, a variety of domains associated with the VQ were also found in proteins from different taxa. These findings reveal an unexpected structural complexity in the VQ protein family and suggest that their biological functions may be more diverse than we currently know.

3.4.2. Phylogenetic and structural analysis of FvVQ proteins

In order to assess the evolutionary relationships of VQ proteins among species, previous studies have used either the full predicted VQ protein sequences (Zhang et al., 2015;Zhou et al., 2016;Wang et al., 2017b;Guo et al., 2018) or only the conserved amino acids of the aligned VQ domain (Kim et al., 2013;Wang et al., 2015b;Chu et al., 2016;Dong et al., 2018) to construct phylogenetic trees. In our study, we have taken into account all this information including the recently identified eight amino acids long upstream motif from the core VQ motif (Zhou et al., 2016), to perform a structure-based sequence alignment of the VQ domains from *F. vesca* in comparisons to *Arabidopsis*, soybean and grapevine by PROMALS3D. These results were subsequently used to construct phylogenetic trees by the N-J method in MEGA, either comparing the four mentioned species (Figure 2 and Figure S2), or for the strawberry FvVQ alone (Figure 3). In both cases, the results show a distribution of the VQ proteins into seven groups, with a group-specific arrangement of the upstream motif sequences (Figure S2 and Figure 3). FvVQ13 is classified alone as outgroup because of its sequence divergence, but with low bootstrap support.

Name	*F.vesca homolog* gene	FANhybrid r1.2 scaffolds	FANhybrid gene id.	*Fragaria x ananassa* RefTrans V1	Domains
FaVQ1	*gene05795-v1.0-hybrid*	FANhyb_rscf00000599.1	*FANhyb_rscf00000599.1.g00006.1*	fananassa_gdr_reftransV1_0027312 *	VQ
				fananassa_gdr_reftransV1_0053740 *	VQ
FaVQ2	*gene12350-v1.0-hybrid*	FANhyb_icon00018020_a.1	*FANhyb_icon00018020_a.1.g00001.1*	no hits found	
FaVQ3	*gene17872-v1.0-hybrid*	FANhyb_icon00002146_a.1	*FANhyb_icon00002146_a.1.g00001.1*	fananassa_gdr_reftransV1_0042573	VQ
				fananassa_gdr_reftransV1_0062571	VQ
FaVQ4	*gene16338-v1.0-hybrid*	FANhyb_rscf00005480.1	*FANhyb_rscf00005480.1.g00001.1*	fananassa_gdr_reftransV1_0054430 *	VQ
FaVQ5	*gene32666-v1.0-hybrid*	FANhyb_icon00017850_a.1	*FANhyb_icon00017850_a.1.g00001.1*	fananassa_gdr_reftransV1_0044247	VQ
FaVQ6	*gene02378-v2.0.a2-hybrid*	FANhyb_icon00041277_a.1	*FANhyb_icon00041277_a.1.g00001.1*	fananassa_gdr_reftransV1_0054841	VQ
		FANhyb_rscf00000298.1	*FANhyb_rscf00000298.1.g00005.1*	fananassa_gdr_reftransV1_0000040	VQ
FaVQ7	*gene19985-v1.0-hybrid*	FANhyb_rscf00000053.1	*FANhyb_rscf00000053.1.g00024.1*	fananassa_gdr_reftransV1_0018796 *	VQ
FaVQ8	*gene02352-v2.0.a2-hybrid*			fananassa_gdr_reftransV1_0019517	VQ
FaVQ9	*gene17096-v2.0.a2-hybrid*	FANhyb_rscf00000298.1	*FANhyb_rscf00000298.1.g00005.1*	fananassa_gdr_reftransV1_0040065 *	VQ
		FANhyb_rscf00000496.1	*FANhyb_rscf00000496.1.g00005.1*	fananassa_gdr_reftransV1_0059358	VQ
FaVQ10	*gene22762-v1.0-hybrid*	FANhyb_rscf00000136.1	*FANhyb_rscf00000136.1.g00007.1*	fananassa_gdr_reftransV1_0068457 *	VQ
FaVQ11	*gene03724-v1.0-hybrid*			fananassa_gdr_reftransV1_0029239	VQ
FaVQ12	*gene25153-v1.0-hybrid*	FANhyb_rscf00000010.1	*FANhyb_rscf00000010.1.g00032.1*	fananassa_gdr_reftransV1_0019677 *	VQ
				fananassa_gdr_reftransV1_0023329	VQ, LRR8, NB-ARC
FaVQ13	*gene26510-v2.0.a2-hybrid*	FANhyb_rscf00006122.1	*FANhyb_rscf00006122.1.g00001.1*	fananassa_gdr_reftransV1_0063702	VQ, LRR8, NB-ARC
				fananassa_gdr_reftransV1_0042571	VQ, NB-ARC
FaVQ14	*gene11973-v1.0-hybrid*	FANhyb_rscf00002726.1	*FANhyb_rscf00002726.1.g00001.1*	fananassa_gdr_reftransV1_0001655 *	VQ
				fananassa_gdr_reftransV1_0058547 *	VQ
FaVQ15	*gene22130-v1.0-hybrid*	FANhyb_rscf00001000.1	*FANhyb_rscf00001000.1.g00002.1*	fananassa_gdr_reftransV1_0059602	VQ
FaVQ16	*gene16601-v1.0-hybrid*	FANhyb_rscf00000324.1	*FANhyb_rscf00000324.1.g00002.1*	fananassa_gdr_reftransV1_0042218	VQ
FaVQ17	*gene00230-v1.0-hybrid*	FANhyb_icon00005101_a.1	*FANhyb_icon00005101_a.1.g00002.1*	fananassa_gdr_reftransV1_0073885	VQ
				fananassa_gdr_reftransV1_0013307	VQ
FaVQ18	*gene21583-v1.0-hybrid*	FANhyb_rscf00004997.1	*FANhyb_rscf00004997.1.g00001.1*	fananassa_gdr_reftransV1_0027413 *	VQ
FaVQ19	*gene18133-v1.0-hybrid*			fananassa_gdr_reftransV1_0033646	VQ
FaVQ20	*gene22640-v1.0-hybrid*	FANhyb_icon00000069_a.1	*FANhyb_icon00000069_a.1.g00001.1*	fananassa_gdr_reftransV1_0038560 *	VQ
FaVQ21	*gene15994-v1.0-hybrid*	FANhyb_icon00002699_a.1	*FANhyb_icon00002699_a.1.g00001.1*	fananassa_gdr_reftransV1_0069556	VQ
FaVQ22	*gene28736-v1.0-hybrid*			no hits found	
FaVQ23	*gene28758-v1.0-hybrid*			no hits found	
FaVQ24	*gene24219-v1.0-hybrid*	FANhyb_icon00016110_a.1	*FANhyb_icon00016110_a.1.g00001.1*	fananassa_gdr_reftransV1_0059293	VQ
FaVQ25	*gene14221-v1.0-hybrid*			fananassa_gdr_reftransV1_0013396	VQ

Table 2. Fragaria x ananassa VQ genes and transcripts homologs to FvVQs. Asterisks indicate transcripts mapped on the FANhybrid r1.2 reference genome by BLAT, with alignments lengths of 95% and 90% identity. BLAT alignments can be found at https://www.rosaceae.org/analysis/230.



Taxa	Taxid	Sequences	Accessions	Domain Names
Streptophyta	35493	4271	pfam05678	VQ
Spermatophyta	58024	59	pfam05678; cl26464	VQ; Atrophin-1
Pentapetalae	1437201	19	pfam05678; cl25764	VQ; PAT1
Pentapetalae	1437201	13	pfam05678; cl27891	VQ; LRR_3
Mesangiospermae	1437183	10	cl26464; pfam05678	Atrophin-1; VQ
rosids	71275	9	pfam05678; cl25887	VQ; RRM
Mesangiospermae	1437183	8	pfam05678; pfam00931 ;cl25992	VQ; NB-ARC ;LRR_8
Brassicaceae	3700	7	pfam07887; pfam05678 ;cl27891	Calmodulin_bind; VQ ;LRR_3
Brassicaceae	3700	5	pfam05678; cl17169	VQ; RRM_SF
Brassica	3705	5	pfam01049; pfam05678	Cadherin_C; VQ
rosids	71275	5	pfam05678; cl28087	VQ; FtsK
Pentapetalae	1437201	4	pfam05678; pfam00931	VQ; NB-ARC
Maleae	721813	3	pfam05678; pfam14723	VQ; SSFA2_C
Oryza sativa	4530	3	cl23735; pfam05678	H4; VQ ;
rosids	71275	3	pfam05678; cl27489	VQ; Glucosaminidase
Brassica	3705	2	pfam05678; pfam11383	VQ; DUF3187
Brassica oleracea var. oleracea	109376	2	pfam05678; cl00031	VQ; ALBUMIN
Embryophyta	3193	2	cl26386; pfam05678	DNA_pol3_gamma3; VQ
Brassica	3705	2	pfam05678; cl27863	VQ; SSL_OB
rosids	71275	2	pfam05678; cl26247	VQ; DNA_pol3_delta2
Erythranthe guttata	4155	2	pfam05678; cl25836	VQ; Retinal
malvids	91836	2	pfam05678; cl25764 ;cl26464	VQ; PAT1 ;Atrophin-1
Arachis	3817	2	pfam05678; cl11531	VQ; ZapB
Pentapetalae	1437201	2	pfam05678; cl26386	VQ; DNA_pol3_gamma3
Populus euphratica	75702	1	pfam05678; pfam13266	VQ; DUF4057

Accesion	Description
XP_015612015.1 GI:1002297180	PREDICTED: putative disease resistance protein RGA3 [Oryza sativa Japonica Group], 1137 aa protein
XP_010673130.1 GI:731324745	PREDICTED: putative disease resistance protein RGA3 [Beta vulgaris subsp. vulgaris], 679 aa protein
BAT07727.1 GI:937932812	Os09g0365000 [Oryza sativa Japonica Group], 1079 aa protein
ALO70146.1 GI:953415386	NBS-LRR-like resistance protein [Oryza sativa], 1137 aa protein
ALO70147.1 GI:953415388	NBS-LRR-like resistance protein [Oryza sativa], 992 aa protein
ALO70148.1 GI:953415390	NBS-LRR-like resistance protein [Oryza sativa], 1107 aa protein
ALO70149.1 GI:953415392	NBS-LRR-like resistance protein [Oryza sativa], 1137 aa protein
XP_021745766.1 GI:1219164272	putative disease resistance protein RGA3 [Chenopodium quinoa], 511 aa protein

Figure S1. Protein domains of strawberry VQ13 found by the Conserved Domain Database (CDD) with tables showing the diversity of VQ protein architectures found along different taxa, and examples of proteins with the same domains found in strawberry VQ13 retrieved by CDART.

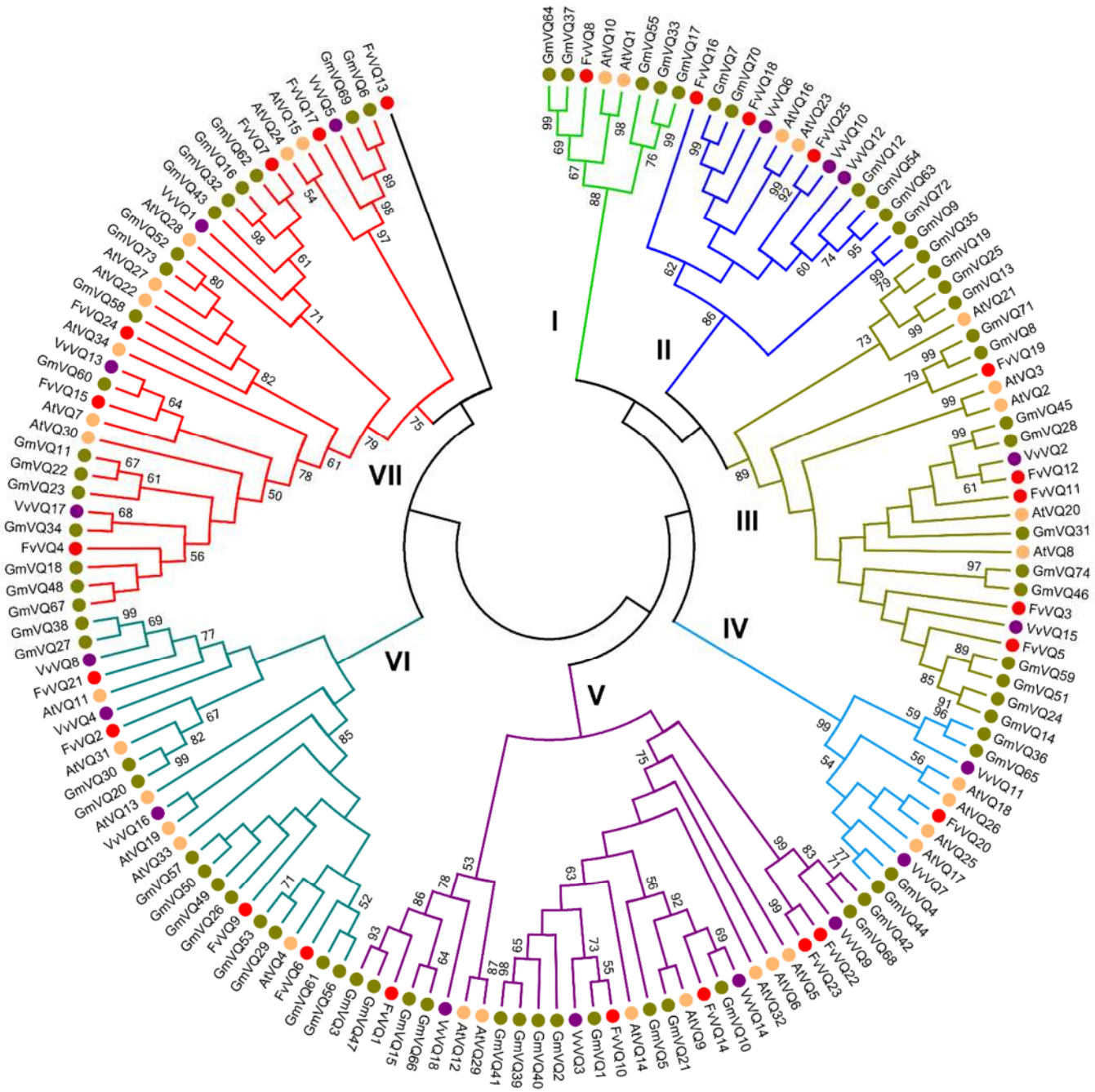


Figure 2. Phylogenetic analysis of strawberry, soybean, grapevine and Arabidopsis VQ proteins. The VQ domain of the four species was aligned by PROMALS3D and an unrooted tree was constructed using MEGA 7.014 by the neighbour-joining method (1000 bootstrap replicates). The groups obtained are sorted by colours and numbered in groups one to seven (I-VII).

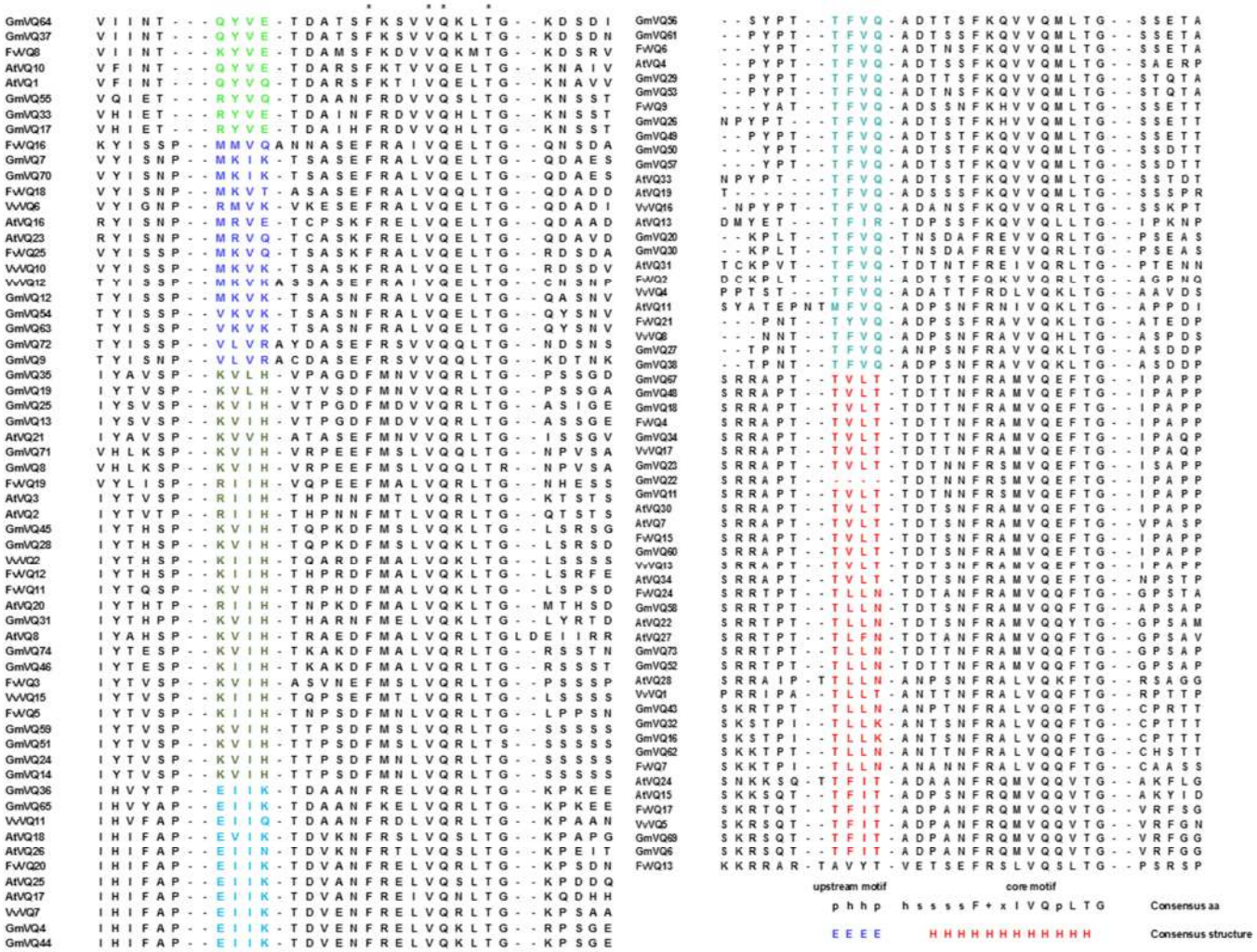


Figure S2. VQ domain alignment corresponding to the phylogenetic tree shown in the main text (Figure2). The same colors to sort the seven different clades obtained and symbols are applicable.

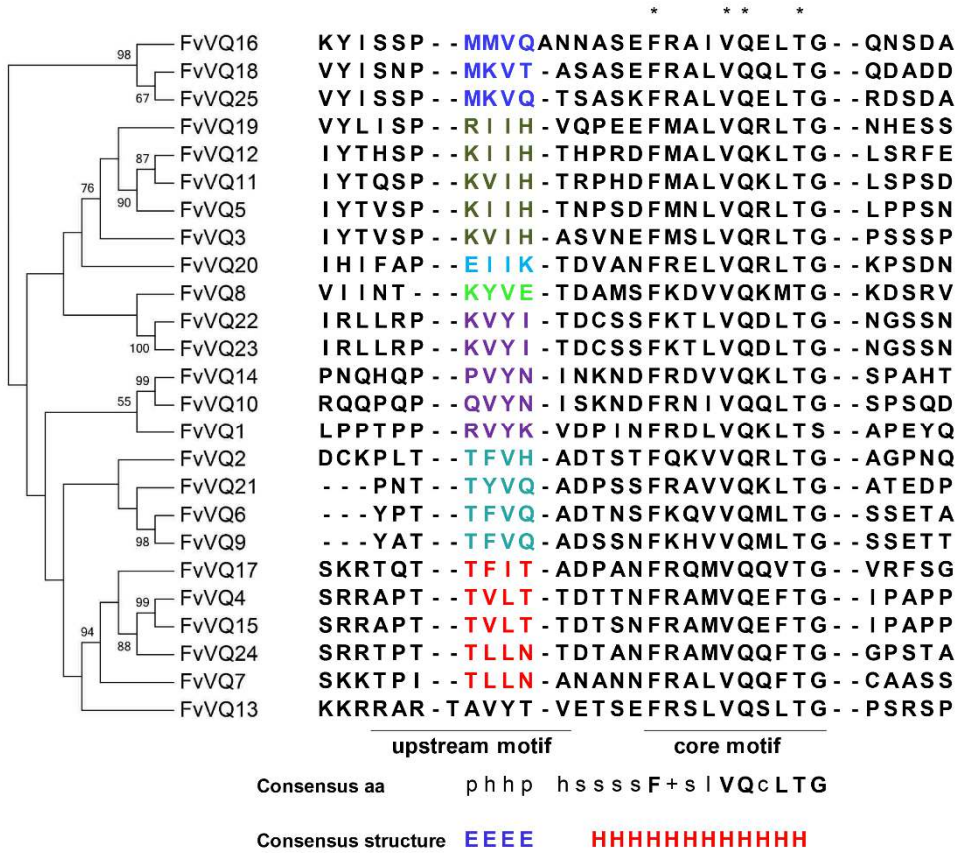


Figure 3. Phylogenetic analysis and multiple sequence alignment of the strawberry VQ protein domain. The VQ domain was aligned by PROMALS3D and an unrooted tree was constructed using MEGA 7.014 by the neighbour-joining method (1000 bootstrap replicates). The seven groups obtained are sorted by the same colours used in the four species tree. The upstream and core motifs of the aligned VQ domain are underlined. Consensus amino acids (aa) and secondary structure are shown. Highly conserved aa are noted by uppercase bold and asterisks. Consensus aa symbols (lowercase): p, polar residues (D, E, H, K, N, Q, R, S, T); h, hydrophobic residues (W, F, Y, M, L, I, V, A, C, T, H); s, small residues (A, G, C, S, V, N, D, T, P); +, positively charged residues (K, R, H); l, aliphatic residues (I, V, L); c, charged (D, E, K, R, H). Consensus structure symbols: E, β-sheet; H, α-helix.

A phylogenetic tree was also constructed using the full protein sequences of the FvVQs and a similar clustering in seven clades was obtained, preserving the group-specific arrangement of the VQ motif sequences (Figure S3), supporting the observation that the VQ domain is the most important determinant for the phylogenetic relationships in the VQ protein family also in strawberry (Zhou et al., 2016).

The structurally-driven sequence alignment reveals the β-sheet-loop-α-helix consensus structure for the VQ domain described by Zhou et al. (2016), who also suggested a critical role of the upstream motif in the WRKY-VQ interaction by modulating the core motif binding specificity, is also preserved in the four species here compared. It is worth to mention that this group distribution does not reflect any kind of intra-group uniformity in terms of exclusivity of binding capabilities to WRKY nor transcriptional regulation activity (activation or repression), at least for those VQs reported in *Arabidopsis*. For example, several VQs placed in different groups interacted with the C-terminal domain of WRKY25 or WRKY33 in YTH (Cheng et al., 2012), while VQs grouped together can exhibit activation, repression or no transcriptional regulation activities in LUC transient expression assays (Li et al., 2014b). This suggest that the VQ motif itself is not the only determinant for the VQ protein functionality, as well as the WRKY binding specificity may be shared by different VQ groups with small differences in their upstream motif sequences.

Besides, complementary structural analysis were carried out in the RaptorX Structure Prediction web server (Kallberg et al., 2012) using the full VQ protein strawberry and Arabidopsis sequences. The results confirmed the consensus secondary structure along the VQ domain, with the notable exception of the upstream motif of FvVQ13, where the β-sheet structure was not predicted. Also, these results, revealed that most of the FvVQ and AtVQ proteins share long protein tracks structurally disordered. Intrinsically disordered proteins (IDPs) are abundant and important components of cellular signalling pathways combining, among others characteristics, the presence of specific recognition motifs, accessible sites for post-translational modifications and a high degree of structural flexibility that allows the interaction with more than one different target (Wright and Dyson, 2015). Most of the biological processes enriched in IDPs found in plants are related to environmental stimuli perception and stress responses, involving protein families like dehydrins, GRAS, NAC and bZIP (Pazos et al., 2013). Consequently, IDP regions could provide new VQ-protein interaction properties to FvVQ and AtVQ proteins making them predictable components of cellular signalling pathways, which can bind different proteins other than WRKY TFs. Thus, MPK3/6 and MPK4 can interact with and phosphorylate respectively, the *Arabidopsis* MVQ subgroup (Pecher et al., 2014) and MSK1 (AtVQ21) (Qiu et al., 2008). Also, AtVQ29 has been shown to bind PIF1 (Li et al., 2014b) and AtVQ32 can bind NDL1 (Cheng et al., 2012). Moreover, *Arabidopsis* VQ12 and VQ29 can bind with themselves or each other to form homo- or heterodimers, as well as with other VQs, being the C-terminal regions responsible for the VQ-VQ interaction (Wang et al., 2015a). The exact domains implicated in these protein-protein associations, if any, remain to be further identified.

Additional motifs included in the FvVQ predicted proteins were analysed by the MEME suite (Figure S3 and Table S3). The conserved VQ domain, containing both the upstream and the core motifs, was clearly recognized in all of the clades, as well as additional motifs featured in the different proteins. Most of such them were also found in the VQ sequences from other species by BLASP searches. Although these motifs potentially indicate function diversity among the VQ proteins, no specific roles are evident or have been reported to date.

3.4.3. Identification of strawberry FvVQ homologs in *Arabidopsis* VQ proteins

In order to infer potential functions of the strawberry VQ proteins, the homology between AtVQs and FvVQs was first investigated using the best reciprocal BLAST hit (BRH) method (Table S4)[1], complemented with the OrthoMCL database and the phylogenetic trees here generated. The result is summarized in Table 3. A consensus among the three methods employed was obtained except for FvVQ5, FvVQ7, FvVQ9 and FvVQ15, either due to the lack of a conclusive best reciprocal AtVQ partner or because the best BLASTP hits did not belong to the same OrthoMCL groups. However, we have assigned putative AtVQ homologs to them taking into account the following considerations. FvVQ3 and AtVQ21 shared the best reciprocal values (bit scores of 65.1 for FvVQ3 as query, and 89.4 for AtVQ21 as query) and FvVQ5 also showed the highest sequence homology values with AtVQ21 but the second best for the reciprocal (bit scores of 83.2 for FvVQ5 as query, and 70.5 for AtVQ21 as query). However, these three proteins clustered together (Figure 2). Therefore, AtVQ21 was also assigned to FvVQ5. In a similar way, AtVQ4 was assigned to FvVQ6 and FvVQ9 (respectively, bit scores of 152 for FvVQ6 as query and 206 for AtVQ4 as query, and of 134 for FvVQ9 as query and 171 for AtVQ4 as query), which also are clustering together. Also, FvVQ7 and FvVQ15 are clustered together within the same clade. However, AtVQ7 was assigned to FvVQ15 as they shared the best reciprocal hits (bit scores of 91.3 for FvVQ15 as query, and 90.9 for AtVQ7 as query). On the contrary, we tentatively assigned AtVQ30 to FvVQ7 only on the basis of the best (non-reciprocal) BLASTP matching pairs (bit scores of 47.4 for FvVQ7 as query) as the best reciprocal hit for AtVQ30 was FvVQ15 (bit scores of 61.6 for AtVQ30 as query), which had already AtVQ7 assigned as the best reciprocal hit option. On the other hand, FvVQ8 and FvVQ16 were classified within group OG5_244916, with no corresponding members of the AtVQ family. We propose AtVQ10 for FvVQ8 as they shared the best reciprocal hits (bit scores of 50.8 for FvVQ8 query, and 51.2 for AtVQ10 as query). Also, FvVQ16 matched the best (non-reciprocal) with AtVQ23 (bit scores of 48.5 for FvVQ16 as query) but AtVQ23 did it both with FvVQ18 as the best hit and FvVQ16 as the best third (bit scores of 55.5 and of 41.2, respectively for AtVQ23 as query). As long as FvVQ16 and FvVQ18 were clustered together and FvVQ18 had already AtVQ16 assigned as its best reciprocal hit (bit scores of 59.3 for FvVQ18 query, and 62.8 for AtVQ16 query),

[1] Esta tabla es demasiado grande para ser representada aquí. Ver publicación.

AtVQ23 was assigned to FvVQ16. For the remaining FvVQs, the OrthoMCL grouping, BRH results and the phylogenetic clustering obtained were coincident. For FvVQ1 two putative orthologs were found, AtVQ12 and AtVQ29, two closely related proteins, but showing the highest homology to AtVQ12. On the other hand, FvVQ13 was classified as member of the group OG5_134032, together with several R-proteins as the RPP13-like protein 1 (At3g14470.1), which was the best BLASTP hit found in the Arabidopsis proteome. Little homology of FvVQ13 with other VQ members was found but restricted to the VQ domain context.

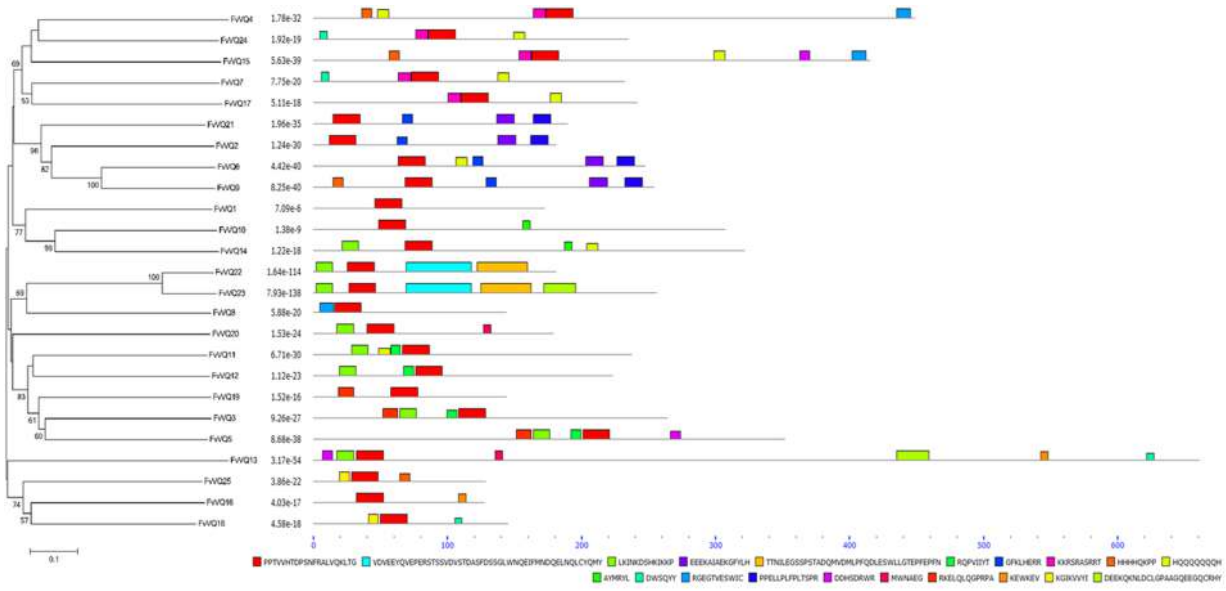


Figure S3. Phylogenetic tree of strawberry VQ proteins and motifs found by MEME. The full FvVQ sequences were aligned by MUSCLE and an unrooted tree was constructed using MEGA 7.014 by the neighbour-joining method (1000 bootstrap replicates). The 20 motifs found are schematized in different colors and the combined mach p-values are shown (see Table S3 for further details).

Motif	Secuence	Width	E-value	Sites	Best possible match
1	PPTVVHTDPSNFRALVQKLTG	21	7.70E-185	25	PPKVYHTDCSNFRALVQKLTG
2	VDVEEYQVEPERSTSSVDVSTDASFDSSGLWNQEIFMNDQELNQLCYQMY	50	4.30E-18	2	VDVEEYQVEPERSTSSVDVSTDASFDSSGLWNQEIFMNDQELNQLCYQMY
3	LKINKDSHKIKKKP	13	7.50E-18	9	LKINKDSHKIKKKP
4	EEEKAIAEKGFYLH	14	9.80E-12	4	EEEKAIAEKGFYLH
5	TTNILEGSSPSTADQMVDMLPFQDLESWLLGTEPFEPFN	39	4.90E-07	2	TTNILEGSSPSTADQMVDMLPFQDLESWLLGTEPFEPFN
6	RQPVIIYT	8	5.30E-03	4	RHPVIIYT
7	GFKLHERR	8	3.60E-02	4	GFKLYERR
8	KKRSRASRRT	10	1.20E+00	5	KKRSRASRRT
9	HHHHQKPP	8	1.50E+00	4	HHHHQKPP
10	HQQQQQQQH	9	1.70E+00	8	HQQQQQQQH
11	AYMRYL	6	3.30E+00	2	AYMRYL
12	DWSQYY	6	2.70E+00	4	DWEQYY
13	RGEGTVESWIC	11	3.30E+00	3	RGEGTVESWIC
14	PPELLPLFPLTSPR	14	5.30E+00	4	EPELLPLFPLTSPR
15	DDHSDRWR	8	5.70E+01	3	DDHDDRWR
16	MWNAEG	6	7.20E+01	2	MWMAEG
17	RKELQLQGPRPA	12	1.50E+02	3	RKELQLQGPRPA
18	KEWKEV	6	3.30E+02	2	KEWKEV
19	KGIKVVYI	8	3.60E+02	2	KGIKVVYI
20	DEEKQKNLDCLGPAAGQEEGQCRHY	25	4.30E+02	2	DEEKVKIVDCLGPHAGQEEGYCRHY

Table S3. Motifs found by MEME in strawberry VQ proteins

Table 3. Orthologs between strawberry and Arabidopsis VQ proteins. The best reciprocal BLASTP hits are marked by asterisks.

OrthoMCL group	AtVQ members	FvVQ protein	Arabidopsis ortholog
OG5_177680	AtVQ12, 29	FvVQ1	AtVQ12 *, AtVQ29
OG5_213230	AtVQ31 (MVQ6)	FvVQ2	AtVQ31 *
OG5_213152	AtVQ21 (MKS1)	FvVQ3	AtVQ21 *
OG5_212399	AtVQ20	FvVQ12	AtVQ20 *
OG5_134032	no AtVQs in this group	FvVQ13	At3g14470.1
OG5_190867	AtVQ9 (MVQ10)	FvVQ14	AtVQ9 *
OG5_164495	AtVQ15,24	FvVQ17	AtVQ24 *
OG5_177741	AtVQ17,25	FvVQ20	AtVQ25*
OG5_147155	AtVQ2,3	FvVQ5	AtVQ21
OG5_170456	AtVQ14 (IKU1, MVQ9)	FvVQ10	AtVQ14 *
NO_GROUP	AtVQ5,13,19,28,30,33	FvVQ9	AtVQ4
OG5_178155	AtVQ34	FvVQ4	AtVQ34
		FvVQ15	AtVQ7*
OG5_150231	AtVQ4 (MVQ1), 11 (MVQ5)	FvVQ6	AtVQ4 *
		FvVQ21	AtVQ11 *
OG5_170861	AtVQ22 (JAV1), 27	FvVQ7	AtVQ30
		FvVQ24	AtVQ22 *
OG5_244916	no AtVQs in this group	FvVQ8	AtVQ10 *
		FvVQ16	AtVQ23
OG5_212106	AtVQ8 (PDE337)	FvVQ11	AtVQ8 *
		FvVQ19	AtVQ8
OG5_189669	AtVQ16 (SIB2), 23 (SIB1)	FvVQ18	AtVQ16 *
		FvVQ25	AtVQ23 *

3.4.4. Predicted *Cis*-control elements within the regulatory region of *FvVQ* genes

To gain insight into the regulation of the strawberry *VQs* gene expression, the putative regulatory regions of *FvVQ* genes were investigated searching for known *cis*-elements. Thus, the 1500bp upstream regions preceding all *FvVQ* coding sequences (except 842bp for *FvVQ10*, 502bp for *FvVQ*17 and 1185bp for *FvVQ*23) were screened in the PLANTCARE database. A wide range of biotic and abiotic stresses, elicitor, phytohormone responsive elements and developmental regulatory sequences were predicted (Table S5)[2]. Interestingly, there are abundance of light, abscisic acid (ABRE), gibberellins (GARE and P-box) and auxin (TGA) responsive elements, which have already been described as important factors in the development and ripening of fruit receptacle and achenes in this non-climacteric fruit (Watson et al., 2002;Csukasi et al., 2011;Symons et al., 2012). Also remarkable is the presence in most of the promoters here analysed, of two elements involved in endosperm gene expression, GCN4- and Skin-1 motifs, suggesting that a number of *FvVQ* genes could be expressed in the endosperm of achenes.

On the other hand, known response elements to salicylic acid (TCA-element) and methyl jasmonate (CGTCA- and TGACG-motifs) were also found in most of the *FvVQs*, suggesting a potential role of these genes in strawberry plant defence. Furthermore, several binding site sequences for important plant transcription factors (TFBS) including NAC, bZIP, bHLH, MYB and WRKY, were found within many *FvVQ* promoter regions (Table S6). Given that WRKY TFs can bind to alternative sequences other than the consensus TTGACY (where Y=C/T), we also searched for WK-boxes (TTTTCCAC), recognized by WRKY TFs carrying the WRKYGKK motif(van Verk et al., 2008) and WT-boxes (YGACTTTT), recognized by the *Arabidopsis* WRKY70 (Machens et al., 2014). Results are summarized in Table S7. One or more W-boxes were found in 23 out of 25 *FvVQs* promoters, with exceptions of *FvVQ21* and *FvVQ25*. No WK-boxes were found, but single WT-box sequences are present in *FvVQ2*, *FvVQ3*, *FvVQ5*, *FvVQ12* and *FvVQ24*. The theoretically-expected frequencies for the consensus W-box and WT-box for both DNA chains within the 1.5-kb promoter sequences of *FvVQs* (GC%: 38.9793 for the FraVesHawaii_1.0 assembly) are roughly 1 per 1.7-kb and 17.4-kb, respectively. Based on this, we consider that the presence of two, or more W-box sequences as well as single WT-boxes within the 1500 bp upstream region of *FvVQ* genes is remarkable and suggests that the expression of some strawberry *VQs* may be regulated by WRKY TFs, including strawberry WRKY70-like TFs.

[2] Estas tablas son demasiado extensas para ser representadas aquí. Consultar publicación.

Table S6. Group promoter analysis of F. vesca VQ genes by PLANTPAN 2.0.

Transcription Factor Binding site with Transcription Factor			
TFBS ID	**Family of TF**	**Support**	**Confidence**
TFmatrixID_0235	Dof	92.0	92.0
TFmatrixID_0445	WRKY	92.0	92.0
TFmatrixID_0382	NAC; NAM	92.0	92.0
TF_motif_seq_0313	(Others)	92.0	92.0
TFmatrixID_0131	AT-Hook	92.0	92.0
TF_motif_seq_0255	AP2; RAV; B3	96.0	96.0
TF_motif_seq_0300	bHLH	96.0	96.0
TF_motif_seq_0244	SBP	96.0	96.0
TF_motif_seq_0410	bHLH	96.0	96.0
TFmatrixID_0227	TCR; CPP	96.0	96.0
TF_motif_seq_0241	ZF-HD	100.0	100.0
TF_motif_seq_0243	GATA; tify	100.0	100.0
TF_motif_seq_0246	Homeodomain; TALE	100.0	100.0
TF_motif_seq_0252	Myb/SANT; MYB; ARR-B	100.0	100.0
TF_motif_seq_0254	AP2; ERF	100.0	100.0
TF_motif_seq_0339	WRKY	100.0	100.0
TF_motif_seq_0257	NF-YB; NF-YA; NF-YC	100.0	100.0
TF_motif_seq_0258	Dehydrin	100.0	100.0
TF_motif_seq_0267	Trihelix	100.0	100.0
TF_motif_seq_0270	WRKY	100.0	100.0
TF_motif_seq_0271	bZIP	100.0	100.0
TF_motif_seq_0302	bHLH	100.0	100.0
TF_motif_seq_0239	Dof	100.0	100.0
TF_motif_seq_0237	GATA; tify	100.0	100.0
TFmatrixID_0193	bZIP	100.0	100.0
TF_motif_seq_0508	SBP	100.0	100.0
Transcription Factor Binding Site only (No association TF)			
TF_motif_seq_0083	D3GMAUX28	92.0	92.0
TF_motif_seq_0455	E2FANTRNR	92.0	92.0
TF_motif_seq_0450	PALINDROMICCBOXGM	92.0	92.0
TF_motif_seq_0315	GT1GMSCAM4	92.0	92.0
TF_motif_seq_0303	CATATGGMSAUR	92.0	92.0
TF_motif_seq_0458	E2FCONSENSUS	96.0	96.0
TF_motif_seq_0260	OSE2ROOTNODULE	96.0	96.0
TF_motif_seq_0375	SEF4MOTIFGM7S	100.0	100.0
TF_motif_seq_0449	AUXRETGA1GMGH3	100.0	100.0
TF_motif_seq_0275	WBOXATNPR1	100.0	100.0
TF_motif_seq_0268	ARR1AT	100.0	100.0
TF_motif_seq_0321	GT1CONSENSUS	100.0	100.0
TF_motif_seq_0265	SORLIP2AT	100.0	100.0
TF_motif_seq_0341	MYB1AT	100.0	100.0
TF_motif_seq_0263	SORLIP1AT	100.0	100.0
TF_motif_seq_0343	ANAERO1CONSENSUS	100.0	100.0
TF_motif_seq_0261	SURECOREATSULTR11	100.0	100.0
TF_motif_seq_0249	ABRELATERD1	100.0	100.0
TF_motif_seq_0248	MYBCOREATCYCB1	100.0	100.0
TF_motif_seq_0009	LS7ATPR1	100.0	100.0

Table S7. WRKY binding sites in FvVQ promoters.

Gene promoter	Motif	Position	Strand	Sequence
FvVQ1p	W-box (5)	1077	-	TTGACT
		1221	-	TTGACC
		1237	+	TTGACC
		279	-	TTGACC
		49	+	TTGACT
FvVQ2p	WT-box	300	+	TGACTTTT
	W-box (4)	1337	-	TTGACT
		1390	-	TTGACT
		1401	-	TTGACT
		470	+	TTGACC
FvVQ3p	WT-box	486	-	TGACTTTT
	W-box (2)	1486	-	TTGACT
		250	-	TTGACT
FvVQ4p	W-box (2)	1300	+	TTGACT
		865	-	TTGACC
FvVQ5p	WT-box	140	+	CGACTTTT
	W-box (2)	367	+	TTGACC
		48	+	TTGACC
FvVQ6p	W-box (4)	1387	-	TTGACC
		172	+	TTGACT
		554	-	TTGACC
		786	+	TTGACC
FvVQ7p	W-box (3)	1405	-	TTGACC
		1430	+	TTGACC
		945	-	TTGACT
FvVQ8p	W-box (6)	1263	+	TTGACC
		1353	-	TTGACC
		230	-	TTGACC
		54	-	TTGACT
		698	+	TTGACC
		874	-	TTGACT
FvVQ9p	W-box (3)	1459	-	TTGACT
		643	+	TTGACT
		985	-	TTGACT
FvVQ10p	W-box (3)	239	+	TTGACC
		276	-	TTGACT
		715	+	TTGACT
FvVQ11p	W-box (4)	1143	+	TTGACC
		1166	-	TTGACC
		215	+	TTGACC
		900	+	TTGACC

Table S7 (continued)

FvVQ12p	WT-box	648	+	TGACTTTT
	W-box	647	+	TTGACT
FvVQ13p	W-box (2)	107	-	TTGACT
		346	+	TTGACC
FvVQ14p	W-box	529	+	TTGACT
FvVQ15p	W-box	965	+	TTGACC
FvVQ16p	W-box	329	+	TTGACT
FvVQ17p	W-box (2)	129	-	TTGACT
		376	-	TTGACT
FvVQ18p	W-box	1471	-	TTGACT
FvVQ19p	W-box	1311	+	TTGACC
FvVQ20p	W-box (2)	721	-	TTGACC
		99	-	TTGACT
FvVQ21p	not found			
FvVQ22p	W-box (3)	1169	+	TTGACC
		1274	-	TTGACC
		86	-	TTGACT
FvVQ23p	W-box (2)	225	+	TTGACC
		334	-	TTGACC
FvVQ24p	WT-box	1332	+	TGACTTTT
	W-box (2)	1338	+	TTGACC
		257	+	TTGACC
FvVQ25p	not found			

3.4.5. Expression profiles of *FaVQ* genes in different tissues and fruit ripening stages

The expression profiles of *VQ*s were analysed in the octoploid strawberry species (cv. Camarosa) by RT-qPCR in different tissues, as well as fruit receptacles and achenes along several ripening stages (Figure 4). Results indicate that *FaVQ1, FaVQ4, FaVQ8, FaVQ9, FaVQ16, FaVQ17, FaVQ24 and FaVQ25* are preferentially expressed in roots, crown, petioles and leaves. Thus, *FaVQ17* and *FaVQ1* were the most expressed genes in root*, FaVQ24* and *FaVQ16* in crown*,* and *FaVQ7, FaVQ8, FaVQ16* and *FaVQ25*, in leaf*, and FaVQ4* and *FaVQ9* both in petiole and flowers, being these two later genes the only ones notably expressed in flowers. Genes *FaVQ7* and *FaVQ16* were not preferentially expressed in fruit but a higher expression in senescent achene. A different set of genes, including *FaVQ2*, *FaVQ3*, *FaVQ10*, *FaVQ11*, *FaVQ12*, *FaVQ13*, *FaVQ14*, *FaVQ15*, *FaVQ19*, and *FaVQ20*, showed preferential expression in fruit, particularly in red ripe achenes. Only the expression of genes *FaVQ6*, *FaVQ10* and *FaVQ21*, was preferentially detected in fruit receptacle and achene, being *FaVQ6* highly expressed in receptacle of small green fruit, *FaVQ10* in red achene and *FaVQ21* highly expressed in over-ripe achene. It is worth to notice that *FaVQ18* was the highest

expressed in senescent achene. It is tentative to propose that the changes observed in the expression pattern of these *FaVQs* during the fruit ripening stages should be associated to changes in the level of hormones along such a fruit process already described. Thus, it has been described that an initial strong increase of auxin occurs, while gibbellerins declines progressively during the ripening process, at the time that abscisic acid (ABA) increases coinciding with the colour development phase (Symons et al., 2012). In addition, auxin and ABA are found at higher levels in achenes than in receptacles, depending of the developmental stage. Further studies are needed to provide a better understanding of putative roles of FaVQs in this processes.

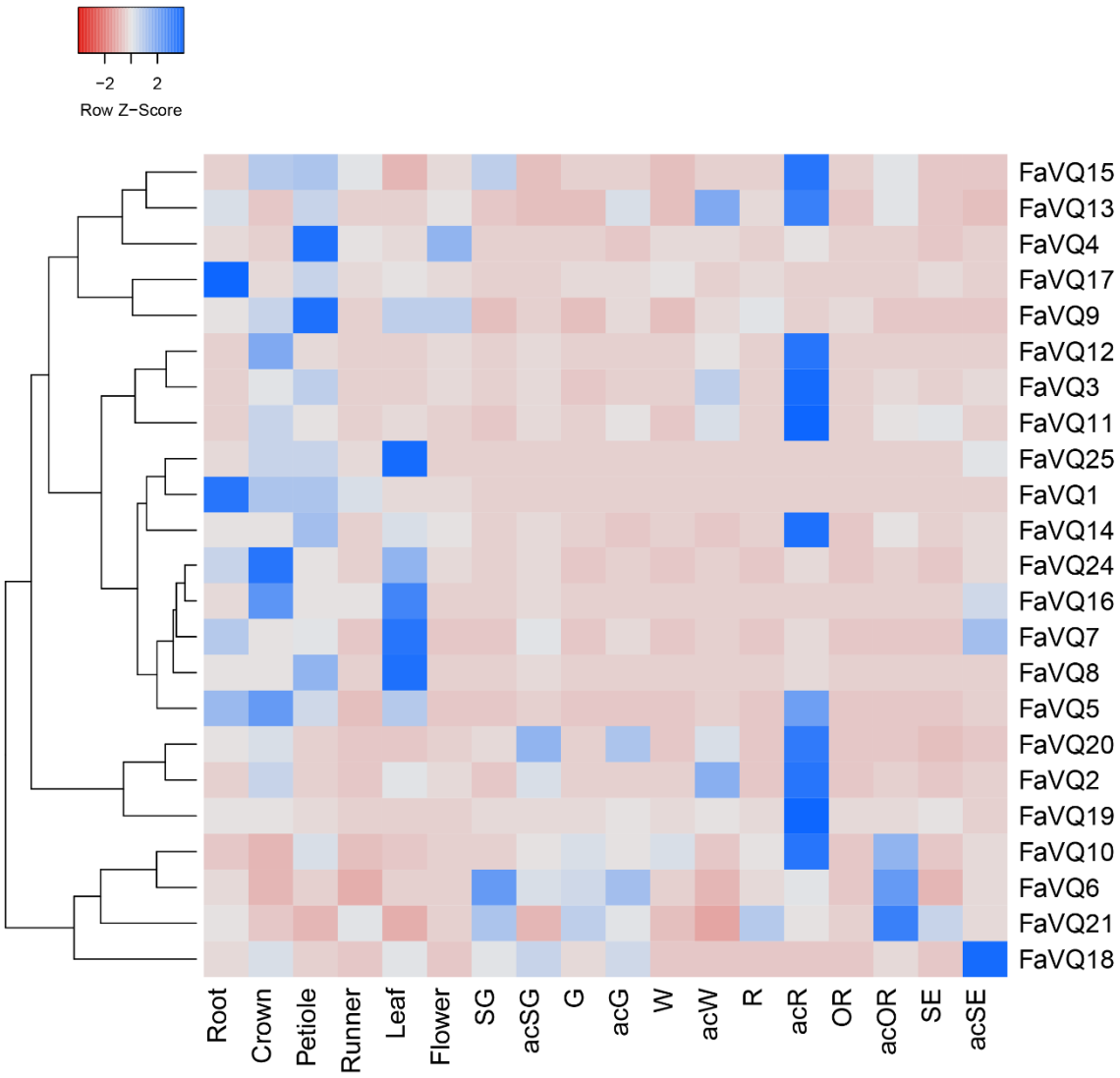


Figure 4. Heatmap representation of the FaVQs expression profiles in different tissues and fruit ripening stages. Abbreviations (fruit receptacles): SG, small green; G, green; W, white; R, red (ripe); OR, overripe; SE, senescent; achenes from the corresponding stages are preceded by an ac- prefix.

3.4.6. Expression analysis of *FaVQ* genes in fruit in response to *C. acutatum* infection.

To explore putative implications of FaVQs in the strawberry defence response against *C. acutatum*, the expression of *FaVQ*s was analysed in fruit of *F. ananassa* cv. Camarosa showing increasing symptoms of anthracnose disease. Since *WRKY33* has been described as a key component in plant defence response to necrotrophic fungi in other plants (Zheng et al., 2006;Birkenbihl et al., 2012) and its interaction with several VQ proteins has been well established (Cheng et al., 2012;Pecher et al., 2014;Jing and Lin, 2015), the expression of *FaWRKY33-1* and *FaWRKY33-2,* two known strawberry *WRKY33* orthologs (Amil-Ruiz et al., 2016) was also studied. Our results showed that the expression of 14 out of the 25 *FaVQs* as well as the two *FaWRKY33s,* was clearly up-regulated in strawberry during pathogen infection but different expression patterns were detected (Figure 5). Thus, a continuous increase of gene expression, which correlates with the infection-induced tissue damage (stages G1 to G3), was found for genes *FaVQ12*, *FaVQ16* and *FaVQ25* (Fig 5a). A second group of genes including *FaVQ1*, *FaVQ5*, *FaVQ7*, *FaVQ9*, *FaVQ11*, *FaVQ19*, *FaVQ21*, *FaVQ24* and *FaWRKY33*.2 showed a maximum accumulation of transcripts at G2 stage but their expressions were reduced at later stages (Fig 5b). On the other hand, the expression pattern of *FaVQ8* and *FavQ17* was very similar and transcripts slightly accumulated at G1 stage, decreasing gene expression at G2 stage before increasing again to a top level at G3 stage (Fig 5c). Contrastingly, *FaWRKY33.1* showed the highest expression at G1 stage and reduced its expression at G2 and G3 stages. The topmost relative expression level (ranging from 90 to 250 times higher than the control uninfected) was for genes *FaVQ1, FaVQ11, FaVQ16* and *FaVQ19,* compared to genes *FaVQ2, FaVQ8, FaVQ12,* and *FaVQ25* (ranging from 20 to 25 times higher than the control uninfected) or genes *FaVQ5, FaVQ9, FaVQ17, FaVQ21,* and *FaVQ24* (ranging from 1.6 to 7 times higher than the control uninfected). Transcription of genes *FaVQ2, FaVQ3, FaVQ4, FaVQ6, FaV10, FaVQ13, FaVQ14, FaVQ15, FaV18, FaVQ20,* remained unchanged, while *FaVQ22* and *FaVQ23* were not detected.

Our results also show that a wide group of *FaVQs* and the two *FaWRKY33* genes were up-regulated after *C. acutatum* infection, and suggest putative implications of these genes in the strawberry defence response. It is well known that AtWRKY33 works as a positive regulator of defence response in Arabidopsis, which can interact with partners controlling its activity. Thus, *MKS1* (*AtVQ21*) and *SIB1* (*AtVQ23) and SIB2 (AtVQ16)* are known regulators of *AtWRKY33* activity and *SIB1, SIB2 and AtWRKY33* showed similar expression patterns and were markedly induced after *B. cinerea* infection (Lai et al., 2011). Accordingly, *FaVQ5* (*MKS1* ortholog), *FaVQ16* and *FaVQ25 (SIB1* orthologs) and the two *FaWRKY33s* were highly up-regulated in strawberry infected fruit while *FaVQ18* (*SIB2 ortholog*) expression did not change significantly.

Recently, a new set of Arabidopsis VQs, have been found to interact with different WRKYs, forming a variety of complexes and being the substrate of MAPKs (Pecher et al., 2014). Such as protein complexes are depending on spatio-temporal VQ and WRKY

expression patterns and defence gene transcription can be modulated by changing the composition of the complexes (Pecher et al., 2014). Thus, VQ8 (MSK1 homolog) was found to interact with WRKY33 as well as with MPK4, suggesting that it may have similar functions to MKS1 (AtVQ21) in defence responses. Curiously, the strawberry VQ8 orthologous genes, *FaVQ11* and *FaVQ19*, were highly induced in fruit after *C. acutatum* infection (Figure 5). It is worth to note that AtVQ8 and the strawberry FaVQ11 and FaVQ19 share high homology with MKS1 (AtVQ21) and FaVQ3 and FaVQ5, respectively, and all of them belong to the same clade/cluster (Figure 2 and 3). Another strawberry WRKY33-interacting VQ orthologous gene, *FaVQ7* (*AtVQ30* ortholog), was also detected to be up-regulated.

The expression of some non WRKY33-interacting VQs was also up-regulated in strawberry in response to *C. acutatum* infection. Thus, the expression of *FaVQ12*, the *VQ20* ortholog acting as negative regulator of the defence responses (Cheng et al., 2012), and *FaVQ17* (*VQ24* ortholog) were drastically increased in G3 infected fruit (Figure 5a, c). Also *FaVQ24*, the JAV1 ortholog acting as negative controller of the JA-mediated plant defence response by interacting with WRKY51 (Hu et al., 2013;Yan et al., 2018), was significantly up-regulated in strawberry (Fig 5b). In addition, *FaVQ1* was also strongly induced in fruit upon *C. acutatum* infection, in agreement with the results described in Arabidopsis(Wang et al., 2015a) and rice (Kim et al., 2013) for its *VQ12* and *VQ29* orthologs, which negatively regulate the defence response against *B. cinerea* in a partially dependent JA-signalling pathway manner. Because *FaVQ1* is also homolog to *VQ29*, which bind WRKY33 in YTH assays (Cheng et al., 2012), interaction with strawberry FaWRKY33 proteins may not be discarded. On the other hand, At*VQ10* (*FaVQ8* ortholog), has been recently described to enhance tolerance to oxidative stress (Luhua et al., 2008), and to participate in the defence response to pathogens (Chen et al., 2018). Thus, At*VQ10* has been shown to interact with *WRKY8* to modulate basal defence against *B. cinerea* (Chen et al., 2018). Curiously, in our study, the strawberry FaVQ8 ortholog was up-regulated after *C. acutatum* infection. In addition, other strawberry WRKYs-interacting VQs orthologous genes like *FaVQ9* (*MVQ1* ortholog) and *FaVQ21* (*MVQ5* ortholog) were also up-regulated. Contrastingly, the expression of some other strawberry WRKY-interacting VQ orthologs, was not altered after pathogen infection. Thus, *FaVQ2* (*MVQ6* ortholog), *FaVQ14* (*MVQ10* ortholog), *FaVQ20* (VQ25 ortholog, which negatively regulates defence response against *P. syringae* but not against *B. cinerea* (Cheng et al., 2012)), *FaVQ15* (*VQ7* ortholog, of unknown function) and *FaVQ3* (another *MKS1* ortholog) did not change their expression in response to *C. acutatum*.

Strawberry *FaVQ* orthologous of *VQ* genes with unknown functions in defence response to date, like *FaVQ4* (*VQ34* ortholog), *FaVQ10* (*IKU1* ortholog) and *FaVQ13* (R protein-VQ) did not change their expression pattern in response to *C. acutatum* infection. Interestingly, the rice gene *OsVQ34*, coding for a FaVQ13 structurally-related protein, was no induced in response to both compatible and incompatible strains of *X. oryzae* (Xoo) suggesting that R protein-VQs may not be regulated in response to pathogens at the transcriptional level (Kim et al., 2013).

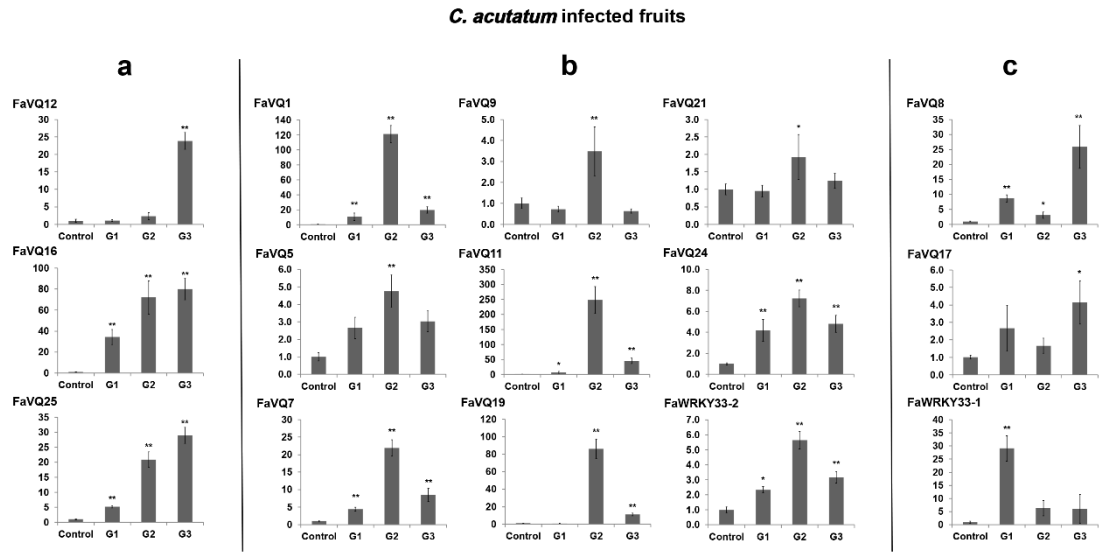


Figure 5. Expression profiles of *FaVQ* genes in anthracnose diseased fruits. The panels (a, b, c) show the different expression patterns described in the main text. Only the genes whose expression were significantly different from the control at any experimental point (Dunnett's test) are shown. Mean, standard error (SE) and significant differences of *P≤0.05 and **P≤0.01 are represented.

3.4.7. Expression analysis of *FaVQ* genes in response to SA and MeJA treatments.

The responsiveness of the *FaVQ* and the two FaWRKY33 genes to the exogenous application of SA and MeJA (the biologically active derivative of jasmonic acid), the two main activators of central defence signalling pathways which regulate responses to pathogens in many plants, was also studied by gene expression analyses. In response to MeJA, the expression of 9 out of the 25 *FaVQs* and of both *FaWRKY33* genes was up-regulated, and 6 other *FaVQ* genes were down-regulated (Figure 6). Thus, the expression pattern of *FaVQ1, FaVQ7, FaVQ8, FaVQ16, FaVQ20, FaVQ25* and *FaWRKY33.2* was similar and a very high increase of transcripts was observed only at 24h after treatment (Figure 6a). A similar expression pattern was also found for genes *FaVQ16* and *FaWRKY33.1* and a continuous and significant increase in gene expression was observed from 6h to 24h after treatment. In addition, transcripts of genes *FaVQ3, FaVQ5, FaVQ11* and *FaVQ24* started to accumulate at 6h after MeJA treatment but their highest expression levels were observed at 12h, diminishing later to lower levels (Figure 6b). Contrastingly, the expression of a third group of genes including *FaVQ12, FaVQ13, FaVQ14* and *FaVQ17* was strongly down-regulated at any times after MeJA treatment (Figure 6c).

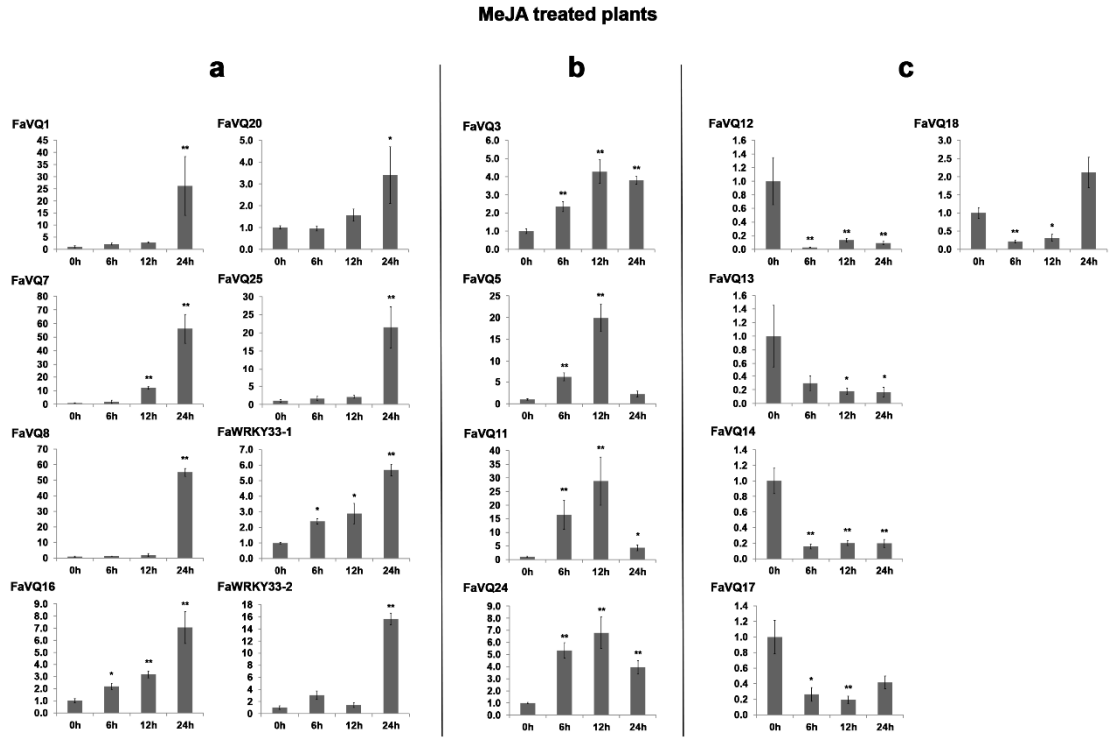


Figure 6. Expression profiles of *FaVQ* genes in MeJA (2 mM) sprayed *in vitro* plants. The panels (a, b, c) show the different expression patterns described in the main text. Only the genes whose expression were significantly different from the control at any experimental point (Dunnett's test) are shown. Mean, standard error (SE) and significant differences of *P≤0.05 and **P≤0.01 are represented.

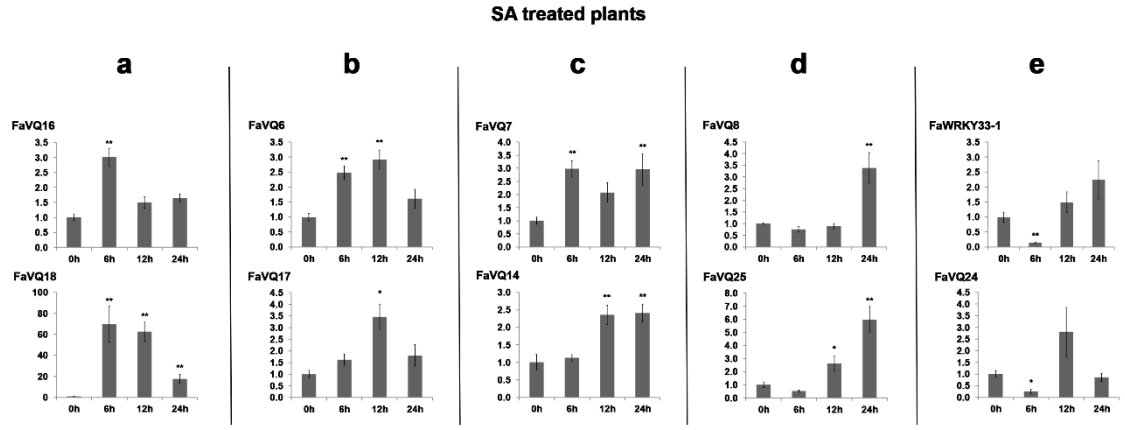


Figure 7. Expression profiles of *FaVQ* genes in SA (5 mM) sprayed *in vitro* plants. The panels (a, b, c, d, e) show the different expression patterns described in the main text. Only the genes whose expression were significantly different from the control at any experimental point (Dunnett's test) are shown. Mean, standard error (SE) and significant differences of *P≤0.05 and **P≤0.01 are represented.

Responsiveness to SA was only detected for 9 out of the 25 *FaVQ* genes tested and the FaWRKY33-1 gene. Thus, the expression of *FaVQ16* and *FaVQ18* was significantly and quickly induced to the highest level at 6h after SA treatment but continuously diminished after that, at later times (Figure 7a). A similar pattern was detected for genes *FaVQ6* and *FaVQ17,* and a very significant increase was only detected at 12 hours (Figure 7b). A much lower but significant up-regulation was also detected at any time for *FaVQ7* and FaVQ14 (Figure 7c) and significant increase was only detected at 24 h after SA treatment for genes *FaVQ8* and *FaVQ25* (Figure 7d). A significant decrease of gene expression was detected at 6h for *FaVQ24* and WRKY33-1 (Figure 7e). No changes in the expression was detected for other *FaVQs* after SA treatment.

Contrastingly, the *FaVQ18* expression (AtVQ16/SIB2 ortholog) was early down-regulated by MeJA but highly induced by SA treatment in strawberry, while SIB1 orthologs FaVQ16 and FaVQ25 were up-regulated by both phytohormones. In addition, *FaVQ7* and *FaVQ8* were highly up-regulated by MeJA but only slightly induced by SA. Accordingly to the results reported in Arabidopsis for JAV1 (Hu et al., 2013), its strawberry *FaVQ24* ortholog was early up-regulated by MeJA. However, SA treatment induced a transient but significant down regulation at 6h. Also, coinciding with the results obtained for VQ12 in Arabidopsis (Wang et al., 2015a), the expression of strawberry *FaVQ1* was up-regulated by MeJA. However, *FaVQ12*, the strawberry ortholog of VQ20 (another known negative regulator of defence (Cheng et al., 2012)), was down-regulated in response to MeJA and not altered by SA treatment. Intriguingly, *FaVQ13* (LRR8-NBS-ARC-VQ protein) expression was highly down-regulated by MeJA treatment, while not responsiveness to SA treatment was detected. On the other hand, FaVQ6 and FaVQ14 were down-regulated by MeJA but significant up-regulation to low levels by SA treatment was detected.

These results are summarized in Figure S4, grouping the *FaVQ* up-regulated genes by the different treatments, as well as the enriched regulatory *cis*-elements found by homology with their *FvVQ* orthologs.

3.4.8. Network interaction analysis of the FaVQ proteins and FaWRKY33 in the response of strawberry to anthracnose disease.

To better understand the complex relationships that the strawberry VQ proteins can establish, we constructed a functional protein association network using STRING 10.5, based on the known interactions of the *Arabidopsis* orthologs uncovered by previous works (Cheng et al., 2012;Pecher et al., 2014;Jing and Lin, 2015;Wang et al., 2015a) and centred in their interactions with WRKY33 TFs. The Figure 8 shows the intricate interactions among FaVQs as positive or negative regulators of the transcriptional activity of FaWRKY33s. A striking fact is that, while FaVQ1 (VQ12 ortholog) seems not able to bind directly to FaWRKY33s, represent a key node for other VQ members that at the same time interact with FaWRKY33s. Accordingly, it has been previously reported that VQ12 have ability to bind to other WRKY33-interacting VQs like MKS1

and its homologs VQ8, VQ10, VQ25 and VQ30, as well as to MPK3/6-targeted VQPs (MVQ1 and MVQ5) (Pecher et al., 2014;Wang et al., 2015a). Notably, VQ12 also interact with JAV1, another JA-responsive negative regulator of defence against pathogens. Altogether, it can be speculated that FaVQ1 (VQ12 ortholog) may operate as an important node for a fine regulation of the JA-mediated response in strawberry by affecting the WRKY33-interacting VQs network. Besides, FaVQ1 could interact specifically with other members of the WRKY TF family, and/or non WRKY33-interacting VQ proteins as JAV1, and have additional roles in regulating gene expression of specific JA-responsive genes involved in the defence against many pathogens. The mechanisms of such regulation, that may imply VQ-VQ protein interactions, would add more complexity to the known models of VQ, WRKY and MPK protein interactions proposed previously (Weyhe et al., 2014).

3.5. Conclusions

A total of 23 VQ encoding genes were confirmed in the genomes of both the wild and the cultivated strawberry using the latest genome annotations and RNAseq. One of the strawberry VQs, Fv/FaVQ13, showed an unusual association with NBS-ARC and LRR8 domains. This new class is named here as R protein-VQ, by analogy with the R protein-WRKY, previously discovered. Also, other VQ proteins with this particular structure have been identified within other species proteomes. None of these proteins have been functionally characterized to date and we only can speculate about their possible roles, but the presence of NBS and LRR domains is expected to be related with pathogen recognition and the regulation of subsequent defence responses.

Strawberry orthologs to main *Arabidopsis* defence-related VQs were found (*MKS1*, *SIB1*, *SIB2*, *JAV1* and *VQ12*, among others), indicating that analogous regulatory mechanisms of defence may exist in these two species. The expression profiles of the *FaVQs* showed tissue- and fruit ripening-dependent patterns. In addition, most of them were regulated in response to anthracnose, as well as to SA and MeJA hormonal treatment, suggesting a role in the strawberry defence responses. These results, lead to consider *FaVQs* as valuable target genes for further functional studies to address breeding programs to improve resistance in this crop.

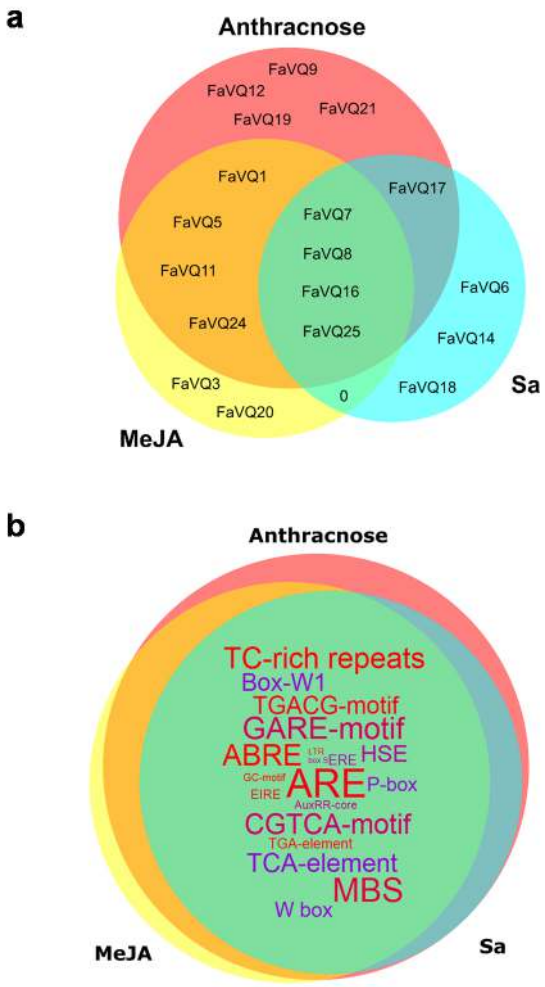


Figure S4. (a) Venn diagrams of up-regulated FaVQs revealing the unique and common genes under the different treatments shown in Fig. 5-7. (b) Venn diagram of the cis-regulatory elements present in the up-regulated FaVQ genes (by homology with their FvVQ orthologs). Only the functional categories "Phytohormone responsiveness" and "Elicitor and stress responses" were used (see Suppl. Table S5). A word cloud was generated, representing the common regulatory sequences (19 out of 27) shared among the up-regulated genes by treatment. Font sizes are proportional to the frequency of the different elements.

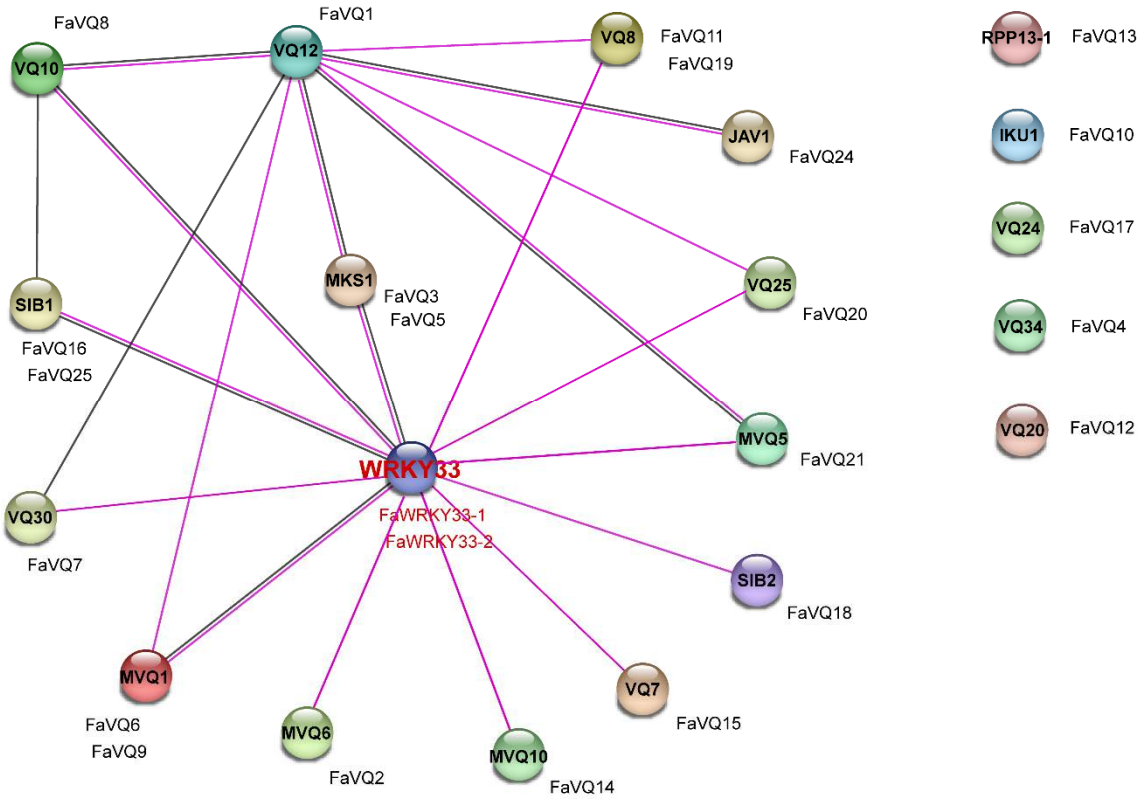


Figure 8. Functional interaction network of FaVQ proteins and FaWRKY33 based in their Arabidopsis orthologs. Nodes are connected by lines indicating experimentally determined interactions (purple) or co-expression (black). Disconnected FaVQ proteins at right side doesn't have known relationships with WRKY33. The original figure was constructed in STRING 10.5 using medium confidence level (0.400) and experiments and co-expression as active interaction sources, then corrected to depict all the known experimental VQ interactions published to date.

3.6. Actualización suplementaria

Tras la realización y publicación de este estudio, se han publicado tanto una nueva anotación del genoma de *F. vesca*, como una nueva secuencia genómica y anotación de *F. x ananassa* (Edger et al., 2019;Li et al., 2019). Por tanto, a continuación se incluye una actualización suplementaria que contiene información adicional sobre la familia de proteínas VQ, tanto en la fresa diploide como en la octoploide.

3.6.1. Métodos

3.6.1.1. Identificación de las proteínas VQ de fresa

Las secuencias de *Fragaria vesca* (anotación v4.0.a2) y *Fragaria x ananassa* cv. Camarosa (Genome Assembly v1.0 & Annotation v1.0.a1) fueron adquiridos de la web The Genome Database for Rosaceae (GDR) (https://www.rosaceae.org/) (Jung et al., 2019). Las secuencias de genes y proteínas anotadas conteniendo el dominio VQ (IPR039829) fueron seleccionadas para su análisis. Se confirmó además que las secuencias candidatas incluyeran dicho dominio, así como dominios adicionales, utilizando la base de datos de dominios conservados (CDD) del NCBI (Marchler-Bauer et al., 2017), seleccionando la base de datos PFAM 31.0.

3.6.1.2. Análisis filogenómicos.

Las secuencias de proteínas VQ de *F. vesca*, *F. x ananassa* y *A. thaliana* fueron alineadas usando MUSCLE. Se construyó un árbol filogenético mediante el método Neighbor-Joining (N-J) en MEGA 7.0 (Kumar et al., 2016). El árbol fue representado y anotado con iTOL (Letunic and Bork, 2019). Se estudió la sintenia comparada entre especies y duplicaciones en genes VQ de fresa en la plataforma web CoGE, utilizando SynMap2 y LAST para encontrar homologías genéticas (Lyons and Freeling, 2008;Haug-Baltzell et al., 2017). Los enlaces persistentes a dichos análisis están recogidos en la Tabla 1 del Capítulo 2. Las tasas de sustitución no sinónimas (Kn) y sinónimas (Ks) entre pares de genes sinténicos se calcularon por codeml, implementado en SynMap2. La ratio Kn/Ks (representada como ω) es usada para estimar la presión de selección. Los valores mayores a 1 indican selección positiva, mientras que los menores de 1 indican selección purificante (negativa) y valores iguales a 1 indican selección neutral (ausencia de evolución) (Yang and Nielsen, 2002).

Para encontrar y clasificar ortólogos de las proteínas VQ de fresa en otras especies, se usó eggNOG-mapper v2 y la eggNOG v5.0 Database (Huerta-Cepas et al., 2016;Huerta-Cepas et al., 2019). Se utilizaron las isoformas proteicas VQ más largas en estos análisis.

3.6.1.3. Otros métodos

Las utilidades contenidas en la plataforma Galaxy (https://usegalaxy.eu/) fueron usadas para la manipulación rutinaria de secuencias y el análisis de los mejores recíprocos mediante BLASTP. La especificidad de los cebadores diseñados previamente fue comprobada sobre la nueva anotación usando la herramienta *Blast search* integrada en el servidor GDR (https://www.rosaceae.org/blast; e-value: 0.01, resto de la configuración por defecto).

3.6.2. Resultados y discusión

3.6.2.1. Actualización de la familia VQ en *F. vesca* y *F. x ananassa*

Los genes y proteínas identificados en *F. vesca* se recogen en la Tabla 4. Su número y propiedades generales resultan ser casi idénticas a las anteriormente descritas, aunque con algunas diferencias, a pesar de que en la nueva anotación se aplicaron los mismos datos derivados de RNA-seq. Esto se debe a que las distintas versiones de secuenciación genética empleadas (v2.0.a2 y v4.0.a2) exhiben pequeñas diferencias entre la composición de sus secuencias y la disposición de las mismas. Así, se observa un cambio en la orientación del segmento que contiene a *FvVQ2*, *FvVQ3* y *FvVQ4* provocando que ahora se hallen en orden inverso (ver Table 1). Por otra parte, los genes anteriormente dudosos, *FvVQ22* y *FvVQ23*, aparecen ahora confirmados en esta nueva anotación y se sitúan en nuevas localizaciones, probablemente como resultado de un ensamblado genómico más completo y exacto. Además, algunos genes y proteínas han experimentado pequeñas modificaciones en su secuencia, aunque la mayoría de las proteínas VQ mantienen una identidad cercana o igual al 100% entre versiones (Tabla 5). Las principales diferencias radican en un incremento en la cantidad de isoformas (variantes alternativas de splicing) encontradas en FvVQ6 y FvVQ13 y en un acortamiento de 55 aa de FvVQ15 (Figura 9), reduciendo su identidad con la versión anterior a algo menos del 81%. Estas diferencias no afectan en ningún caso al dominio VQ, ni a los principales hallazgos ya descritos.

Por otra parte, el recientemente publicado genoma de *F. x ananassa* cv. Camarosa nos ha permitido identificar 96 ortólogos en la fresa octoploide, con una identidad promedio de 97,74% con respecto a los genes de *F. vesca* y, en general, conservando una alta colinealidad (Tabla 6). Los genes *FaVQ* fueron nombrados por tres criterios: primero, por su homología con las *FvVQ* (*FaVQ1* a *FaVQ25*), obtenida por LAST; seguido por una letra indicando el donante del subgenoma (A, *F. nipponica*; B, *F. iinumae*; C, *F. viridis*; D, *F.vesca*) (Edger et al., 2019); por último, numerando las duplicaciones génicas, si las hubiera (.1, .2, etc.).

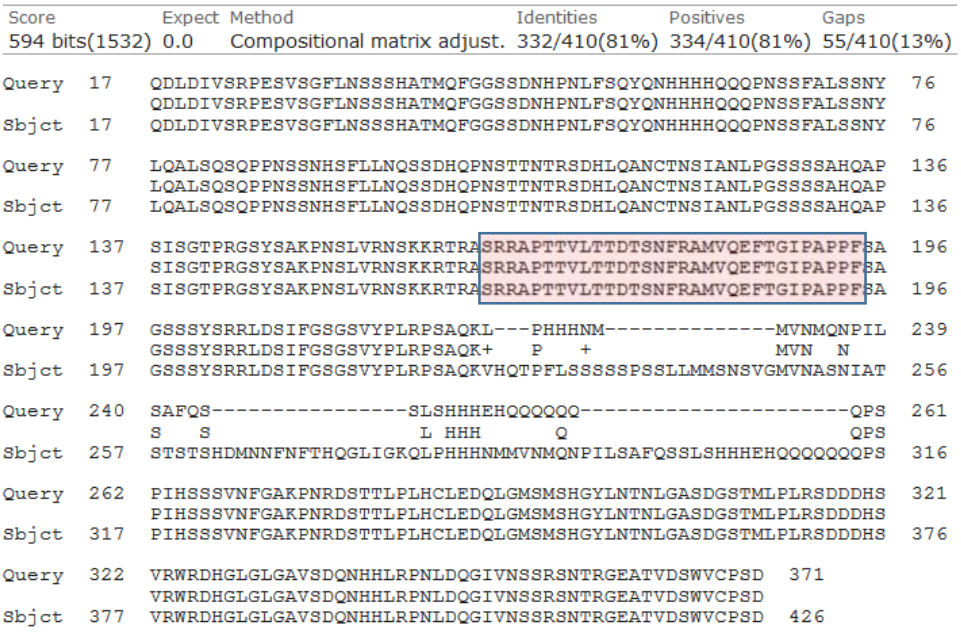
Tabla 4. Descripción de la familia VQ en *F. vesca*

Nombre	Gen Fv	Localización genómica	Isoformas	Longitud (aa)
FvVQ1	*FvH4_1g18020*	Fvb1_v4.0.a1:10489644..10490177 (+ strand)		177
FvVQ2	*FvH4_1g25800*	Fvb1_v4.0.a1:17718022..17718582 (+ strand)		460
FvVQ3	*FvH4_1g24040*	Fvb1_v4.0.a1:15898605..15899450 (- strand)		281
FvVQ4	*FvH4_1g21110*	Fvb1_v4.0.a1:13142222..13143604 (- strand)		186
FvVQ5	*FvH4_2g14340*	Fvb2_v4.0.a1:12606946..12607602 (+ strand)		218
FvVQ6	*FvH4_3g09030*	Fvb3_v4.0.a1:5273026..5274130 (- strand)	4	255
FvVQ7	*FvH4_3g12720*	Fvb3_v4.0.a1:7614861..7615577 (+ strand)		238
FvVQ8	*FvH4_3g16620*	Fvb3_v4.0.a1:10474252..10479279 (+ strand)		125
FvVQ9	*FvH4_3g19800*	Fvb3_v4.0.a1:12956668..12957453 (+ strand)		261
FvVQ10	*FvH4_4g21960*	Fvb4_v4.0.a1:24915439..24916386 (+ strand)		315
FvVQ11	*FvH4_4g27220*	Fvb4_v4.0.a1:28308147..28308881 (- strand)		244
FvVQ12	*FvH4_5g11980*	Fvb5_v4.0.a1:6775531..6776220 (- strand)		229
FvVQ13	*FvH4_5g34700*	Fvb5_v4.0.a1:25409068..25412073 (+ strand)	7	678
FvVQ14	*FvH4_5g36700*	Fvb5_v4.0.a1:26920191..26921183 (+ strand)		330
FvVQ15	*FvH4_5g37690*	Fvb5_v4.0.a1:27740383..27741663 (- strand)		371
FvVQ16	*FvH4_6g02470*	Fvb6_v4.0.a1:1421666..1422061 (+ strand)		131
FvVQ17	*FvH4_6g07790*	Fvb6_v4.0.a1:4662546..4663292 (+ strand)		248
FvVQ18	*FvH4_6g25410*	Fvb6_v4.0.a1:19287384..19287833 (+ strand)		149
FvVQ19	*FvH4_6g21030*	Fvb6_v4.0.a1:14564918..14565364 (+ strand)		148
FvVQ20	*FvH4_6g22390*	Fvb6_v4.0.a1:16062707..16063261 (- strand)		184
FvVQ21	*FvH4_6g29370*	Fvb6_v4.0.a1:22652710..22653297 (+ strand)		195
FvVQ22	*FvH4_6g38531*	Fvb6_v4.0.a1:30488778..30489825 (+ strand)		177
FvVQ23	*FvH4_6g38710*	Fvb6_v4.0.a1:30621661..30625091 (- strand)		263
FvVQ24	*FvH4_6g39500*	Fvb6_v4.0.a1:31182931..31183656 (+ strand)		241
FvVQ25	*FvH4_7g03520*	Fvb7_v4.0.a1:3885876..3886274 (+ strand)		132

Tabla 5. Comparativa de las proteínas VQ entre las dos anotaciones usadas en este estudio. En caso de existir varias isoformas, se usaron las de mayor longitud.

v4.0.a2	Longitud (aa)	v2.0.a2	Longitud (aa)	Identidad (%)
FvVQ1	177	FvVQ1	177	100
FvVQ2	186	FvVQ2	186	100
FvVQ3	281	FvVQ3	271	100
FvVQ4	460	FvVQ4	460	100
FvVQ5	218	FvVQ5	361	100
FvVQ6	255	FvVQ6	254	99.608
FvVQ7	238	FvVQ7	238	100
FvVQ8	125	FvVQ8	125	100
FvVQ9	261	FvVQ9	261	100
FvVQ10	315	FvVQ10	315	100
FvVQ11	244	FvVQ11	244	100
FvVQ12	229	FvVQ12	229	100
FvVQ13	678	FvVQ13	678	100
FvVQ14	330	FvVQ14	330	100
FvVQ15	371	FvVQ15	426	80.976
FvVQ16	131	FvVQ16	131	100
FvVQ17	248	FvVQ17	248	100
FvVQ18	149	FvVQ18	149	100
FvVQ19	148	FvVQ19	148	100
FvVQ20	184	FvVQ20	184	100
FvVQ21	195	FvVQ21	195	100
FvVQ22	177	FvVQ22	177	98.305
FvVQ23	263	FvVQ23	263	100
FvVQ24	241	FvVQ24	241	100
FvVQ25	132	FvVQ25	132	100

En cuanto a la especifidad de unión de los cebadores diseñados para RT-qPCR, en *F. vesca* y *F. x ananassa* no se detectaron cambios significativos en las secuencias diana de los genes correspondientes de ambas especies, aunque la presencia de polimorfismos en algunos genes homoeólogos puede reducir la especificidad de unión de algunos cebadores a sus secuencias diana (datos no mostrados). Por otra parte, en *F. x ananassa* cv. Camarosa se comprobó que los cebadores diseñados para *VQ22* y *VQ23* sólo son válidos para los homoeólogos *FaVQ22D* y *FaVQ23B-23D*, respectivamente. Sin embargo, recordamos que no se consiguieron amplicones previamente usando estos cebadores, ni en la especie diploide ni en la octoploide, usando como molde ADN genómico o ADNc.



Score	Expect	Method	Identities	Positives	Gaps
594 bits(1532)	0.0	Compositional matrix adjust.	332/410(81%)	334/410(81%)	55/410(13%)

```
Query  17   QDLDIVSRPESVSGFLNSSSHATMQFGGSSDNHPNLFSQYQNHHHHQQQPNSSFALSSNY  76
            QDLDIVSRPESVSGFLNSSSHATMQFGGSSDNHPNLFSQYQNHHHHQQQPNSSFALSSNY
Sbjct  17   QDLDIVSRPESVSGFLNSSSHATMQFGGSSDNHPNLFSQYQNHHHHQQQPNSSFALSSNY  76

Query  77   LQALSQSQPPNSSNHSFLLNQSSDHQPNSTTNTRSDHLQANCTNSIANLPGSSSSAHQAP  136
            LQALSQSQPPNSSNHSFLLNQSSDHQPNSTTNTRSDHLQANCTNSIANLPGSSSSAHQAP
Sbjct  77   LQALSQSQPPNSSNHSFLLNQSSDHQPNSTTNTRSDHLQANCTNSIANLPGSSSSAHQAP  136

Query  137  SISGTPRGSYSAKPNSLVRNSKKRTRASRRAPTTVLTTDTSNFRAMVQEFTGIPAPPFSA  196
            SISGTPRGSYSAKPNSLVRNSKKRTRASRRAPTTVLTTDTSNFRAMVQEFTGIPAPPFSA
Sbjct  137  SISGTPRGSYSAKPNSLVRNSKKRTRASRRAPTTVLTTDTSNFRAMVQEFTGIPAPPFSA  196

Query  197  GSSSYSRRLDSIFGSGSVYPLRPSAQKL---PHHHNM--------------MVNMQNPIL  239
            GSSSYSRRLDSIFGSGSVYPLRPSAQK+   P   +                MVN  N
Sbjct  197  GSSSYSRRLDSIFGSGSVYPLRPSAQKVHQTPFLSSSSSPSSLLMMSNSVGMVNASNIAT  256

Query  240  SAFQS----------------SLSHHHEHQQQQQQ----------------------QPS  261
            S   S                  L HHH      Q                      QPS
Sbjct  257  STSTSHDMNNFNFTHQGLIGKQLPHHHNMMVNMQNPILSAFQSSLSHHHEHQQQQQQQPS  316

Query  262  PIHSSSVNFGAKPNRDSTTLPLHCLEDQLGMSMSHGYLNTNLGASDGSTMLPLRSDDDHS  321
            PIHSSSVNFGAKPNRDSTTLPLHCLEDQLGMSMSHGYLNTNLGASDGSTMLPLRSDDDHS
Sbjct  317  PIHSSSVNFGAKPNRDSTTLPLHCLEDQLGMSMSHGYLNTNLGASDGSTMLPLRSDDDHS  376

Query  322  VRWRDHGLGLGAVSDQNHHLRPNLDQGIVNSSRSNTRGEATVDSWVCPSD  371
            VRWRDHGLGLGAVSDQNHHLRPNLDQGIVNSSRSNTRGEATVDSWVCPSD
Sbjct  377  VRWRDHGLGLGAVSDQNHHLRPNLDQGIVNSSRSNTRGEATVDSWVCPSD  426
```

Figura 9. Alineamiento, mediante BLASTP, de las secuencias de FvVQ15 correspondientes a las versiones v4.0.a2 (query) y v2.0.a2 (subject). Se han recuadrado el dominio VQ y las secuencias adyacentes.

En general, los genes VQ de la fresa diploide y octoploide están muy conservados y mantienen sintenia y colinealidad, compartiendo una alta identidad entre sí a la vez que una baja tasa de sustitución de nucleótidos en sus secuencias, tanto sinónimas como no sinónimas, mostrando la mayoría signos de estar bajo presión de selección negativa ($\omega < 1$), indicando que la presión de selección natural se ejerce hacia la conservación de las secuencias y, posiblemente, de las funciones (Tabla 6). Encontramos, sin embargo, algunas excepciones (*FaVQ12D.1*, *FaVQ13D* y *FaVQ16A*) en las que $\omega > 1$, indicando que se están seleccionando mutaciones que pudieran presentar alguna característica ventajosa para la planta. En el caso de *F. x ananassa* también se detecta la pérdida de algunos genes homoeólogos (por ejemplo, *FaVQ2B* y *FaVQ2D*) y algunas duplicaciones génicas no encontradas en *F. vesca* (por ejemplo, *FaVQ11D.1* y *FaVQ11D.2*). Además, observamos pérdidas de colinealidad entre algunos genes ortólogos FvVQ y FaVQ, entre los que se cuentan aquellos ortólogos de los parálogos *FvVQ4-FvVQ15, FvVQ7-FvVQ24* y *FvVQ11-FvVQ12*, que no conservan la colinealidad en ninguno de los subgenomas de *F. x ananassa* (ver Tabla 8 más adelante). Esto indica que el genoma octoploide ha sufrido intensas reorganizaciones genéticas en aquellas regiones cromosómicas que contienen a los genes VQ, tal como se encontró en el estudio de la familia *FaWRKY* (Capítulo 2).

El mapa cromosómico de la familia VQ en la fresa diploide y octoploide está representado en la Figura 10, permitiendo una mejor comparación entre la organización genómica de los genes VQ en ambas especies y entre los subgenomas que componen el genoma octoploide.

Tabla 6. Genes VQ ortólogos de *F. vesca* y *F. x ananassa*

Gen Fv	Ortólogo en Fxa		Identidad (%)	Kn	Ks	ω
FvVQ1	*FaVQ1A*	*augustus_masked-Fvb1-3-processed-gene-104.3-mRNA-1*	98.13	0.0085	0.0247	0.3441
	FaVQ1B	*augustus_masked-Fvb1-2-processed-gene-116.2-mRNA-1*	97.19	0.017	0.0384	0.4427
	FaVQ1C	*augustus_masked-Fvb1-1-processed-gene-175.1-mRNA-1*	97.87	0.0355	0.0483	0.7350
	FaVQ1D	*augustus_masked-Fvb1-4-processed-gene-93.11-mRNA-1*	99.63	0	0.0107	NA
FvVQ2	*FaVQ2A*	*snap_masked-Fvb1-3-processed-gene-191.9-mRNA-1*	97.86	0.0201	0.028	0.7179
	FaVQ2C	*augustus_masked-Fvb1-1-processed-gene-98.3-mRNA-1*	97.08	0.0196	0.075	0.2613
FvVQ3	*FaVQ3A*	*augustus_masked-Fvb1-3-processed-gene-163.11-mRNA-1*	97.97			
	FaVQ3B	*augustus_masked-Fvb1-2-processed-gene-175.10-mRNA-1*	97.54			
	FaVQ3C	*augustus_masked-Fvb1-1-processed-gene-113.9-mRNA-1*	96.68			
	FaVQ3D	*augustus_masked-Fvb1-4-processed-gene-142.1-mRNA-1*	98.99			
FvVQ4	*FaVQ4A*	*augustus_masked-Fvb1-3-processed-gene-134.5-mRNA-1*	96.42	0.0196	0.0651	0.3011
	FaVQ4B	*augustus_masked-Fvb1-2-processed-gene-142.1-mRNA-1*	97.28	0.0126	0.0632	0.1994
	FaVQ4C	*snap_masked-Fvb1-1-processed-gene-149.22-mRNA-1*	96.25	0.0115	0.0573	0.2007
	FaVQ4D	*augustus_masked-Fvb1-4-processed-gene-119.0-mRNA-1*	99.1	0.0027	0.0121	0.2231
FvVQ5	*FaVQ5A*	*augustus_masked-Fvb2-1-processed-gene-84.5-mRNA-1*	98.33	0.0082	0.026	0.3154
	FaVQ5B	*augustus_masked-Fvb2-4-processed-gene-108.0-mRNA-1*	98.02	0.0156	0.0163	0.9571
	FaVQ5C	*augustus_masked-Fvb2-3-processed-gene-128.10-mRNA-1*	98.33	0.0108	0.0173	0.6243
	FaVQ5D	*augustus_masked-Fvb2-2-processed-gene-120.8-mRNA-1*	99.7	0.0019	0.008	0.2375
FvVQ6	*FaVQ6A*	*augustus_masked-Fvb3-3-processed-gene-32.0-mRNA-1*	98.97	1.6278	2.4331	0.6690
	FaVQ6B	*augustus_masked-Fvb3-2-processed-gene-48.12-mRNA-1*	98.66	1.5379	4.0626	0.3786
	FaVQ6C	*augustus_masked-Fvb3-1-processed-gene-270.10-mRNA-1*	97.68	0.0227	0.0518	0.4382
	FaVQ6D	*augustus_masked-Fvb3-4-processed-gene-246.0-mRNA-1*	99.87	0.0153	0.028	0.5464
FvVQ7	*FaVQ7A*	*snap_masked-Fvb3-3-processed-gene-56.14-mRNA-1*	96.65	0.0122	0.0849	0.1437
	FaVQ7B	*augustus_masked-Fvb3-2-processed-gene-68.3-mRNA-1*	94.31	0.0168	0.1437	0.1169
	FaVQ7C	*augustus_masked-Fvb3-1-processed-gene-250.10-mRNA-1*	95.28	0.0218	0.1063	0.2051
	FaVQ7D	*augustus_masked-Fvb3-4-processed-gene-233.7-mRNA-1*	99.86	0	0.0041	NA
FvVQ8	*FaVQ8A*	*augustus_masked-Fvb3-3-processed-gene-78.13-mRNA-1*	96.94	0.0496	0.2233	0.2221
	FaVQ8B	*augustus_masked-Fvb3-2-processed-gene-93.4-mRNA-1*	98.47	0.0388	0.1731	0.2241
	FaVQ8C	*snap_masked-Fvb3-1-processed-gene-225.24-mRNA-1*	97.55	0.0338	0.1496	0.2259
	FaVQ8D	*snap_masked-Fvb3-4-processed-gene-211.33-mRNA-1*	99.08	0.0358	0.1332	0.2688
FvVQ9	*FaVQ9A*	*augustus_masked-Fvb3-3-processed-gene-99.11-mRNA-1*	96.16	0.0122	0.0457	0.2670
	FaVQ9B	*augustus_masked-Fvb3-2-processed-gene-107.2-mRNA-1*	96.16	0.0071	0.0436	0.1628
	FaVQ9C	*augustus_masked-Fvb3-1-processed-gene-203.9-mRNA-1*	96.54	0.0051	0.0423	0.1206
	FaVQ9D	*augustus_masked-Fvb3-4-processed-gene-190.5-mRNA-1*	99.75	0.0019	0.0042	0.4524
FvVQ10	*FaVQ10B*	*snap_masked-Fvb4-4-processed-gene-82.28-mRNA-1*	98.52	0.0062	0.0457	0.1357
FvVQ11	*FaVQ11A*	*snap_masked-Fvb4-2-processed-gene-47.30-mRNA-1*	96.05	0.0036	0.018	0.2000
	FaVQ11B	*augustus_masked-Fvb4-4-processed-gene-50.7-mRNA-1*	95.53	0.0272	0.039	0.6974
	FaVQ11C	*augustus_masked-Fvb4-1-processed-gene-149.10-mRNA-1*	96.61	0.029	0.038	0.7632
	FaVQ11D.1	*snap_masked-Fvb4-3-processed-gene-53.38-mRNA-1*	99.32			
	FaVQ11D.2	*snap_masked-Fvb4-3-processed-gene-81.24-mRNA-1*	99.32	0.0036	0.018	0.2000
FvVQ12	*FaVQ12A*	*augustus_masked-Fvb5-4-processed-gene-60.11-mRNA-1*	94.66			
	FaVQ12B	*augustus_masked-Fvb5-3-processed-gene-213.6-mRNA-1*	94.83	0.0191	0.0277	0.6895
	FaVQ12C	*augustus_masked-Fvb5-2-processed-gene-68.17-mRNA-1*	96.23			
	FaVQ12D.1	*augustus_masked-Fvb5-1-processed-gene-61.12-mRNA-1*	99.13	0.0045	0.0044	1.0227
	FaVQ12D.2	*augustus_masked-Fvb5-1-processed-gene-77.12-mRNA-1*	99.71	0.0023	0.0044	0.5227
FvVQ13	*FaVQ13C*	*maker-Fvb5-2-snap-gene-213.46-mRNA-1*	96.37	0.5968	2.5436	0.2346
	FaVQ13D	*maker-Fvb5-1-augustus-gene-263.34-mRNA-1*	99.71	0.0035	0.0017	2.0588

Tabla 6 (continuación)

FvVQ14	*FaVQ14A*	*augustus_masked-Fvb5-4-processed-gene-226.6-mRNA-1*	95.98	0.0142	0.0692	0.2052
	FaVQ14B	*augustus_masked-Fvb5-3-processed-gene-28.26-mRNA-1*	95.63	0.0129	0.0806	0.1600
	FaVQ14C	*augustus_masked-Fvb5-2-processed-gene-227.1-mRNA-1*	96.88	0.0159	0.0624	0.2548
	FaVQ14D	*augustus_masked-Fvb5-1-processed-gene-277.3-mRNA-1*	99.09	0.0084	0.0115	0.7304
FvVQ15	*FaVQ15A*	*augustus_masked-Fvb5-4-processed-gene-233.14-mRNA-1*	98.06	0.0177	0.0275	0.6436
	FaVQ15B	*augustus_masked-Fvb5-3-processed-gene-19.13-mRNA-1*	97.35	0.0152	0.0413	0.3680
	FaVQ15C	*augustus_masked-Fvb5-2-processed-gene-234.6-mRNA-1*	98.21	0.0141	0.0166	0.8494
	FaVQ15D	*augustus_masked-Fvb5-1-processed-gene-283.9-mRNA-1*	100	0	0.0106	NA
FvVQ16	*FaVQ16A*	*augustus_masked-Fvb6-2-processed-gene-260.1-mRNA-1*	98.73	0.016	0.0081	1.9753
	FaVQ16B.1	*augustus_masked-Fvb6-3-processed-gene-400.10-mRNA-1*	98.23			
	FaVQ16B.2	*augustus_masked-Fvb6-3-processed-gene-404.14-mRNA-1*	98.48	0.014	0.0213	0.6573
	FaVQ16C	*augustus_masked-Fvb6-4-processed-gene-34.3-mRNA-1*	99.24			
	FaVQ16D.1	*augustus_masked-Fvb6-1-processed-gene-346.18-mRNA-1*	99.75	0.0039	0	NA
	FaVQ16D.2	*augustus_masked-Fvb6-1-processed-gene-348.13-mRNA-1*	99.75	0	0.0081	NA
FvVQ17	*FaVQ17A*	*maker-Fvb6-2-augustus-gene-231.24-mRNA-1*	96.25			
	FaVQ17C	*maker-Fvb6-4-snap-gene-56.48-mRNA-1*	96.34			
	FaVQ17D	*maker-Fvb6-1-snap-gene-316.93-mRNA-1*	99.6			
FvVQ18	*FaVQ18A*	*augustus_masked-Fvb6-2-processed-gene-100.5-mRNA-1*	97.57	0.003	0.0673	0.0446
	FaVQ18B	*augustus_masked-Fvb6-3-processed-gene-188.4-mRNA-1*	98.23	0.0029	0.0441	0.0658
	FaVQ18C	*augustus_masked-Fvb6-4-processed-gene-185.3-mRNA-1*	97.79	0.0088	0.0409	0.2152
FvVQ19	*FaVQ19A*	*augustus_masked-Fvb6-2-processed-gene-134.6-mRNA-1*	97.45	0.0423	0.0551	0.7677
	FaVQ19B.1	*snap_masked-Fvb6-3-processed-gene-229.17-mRNA-1*	98.43	0.0118	0.032	0.3688
	FaVQ19B.2	*snap_masked-Fvb6-3-processed-gene-248.14-mRNA-1*	98.43			
	FaVQ19C	*augustus_masked-Fvb6-4-processed-gene-152.6-mRNA-1*	97.99	0.0063	0.037	0.1703
	FaVQ19D	*augustus_masked-Fvb6-1-processed-gene-218.2-mRNA-1*	100	0	0	NA
FvVQ20	*FaVQ20A*	*augustus_masked-Fvb6-2-processed-gene-124.0-mRNA-1*	97.99	0.0121	0.0678	0.1785
	FaVQ20B	*augustus_masked-Fvb6-3-processed-gene-214.7-mRNA-1*	97.03	1.3586	2.569	0.5288
	FaVQ20C	*augustus_masked-Fvb6-4-processed-gene-162.8-mRNA-1*	97.82	0.0144	0.0433	0.3326
	FaVQ20D.1	*augustus_masked-Fvb6-1-processed-gene-200.6-mRNA-1*	99.82	0.0023	0	NA
	FaVQ20D.2	*augustus_masked-Fvb6-1-processed-gene-201.9-mRNA-1*	99.82	0.0023	0	NA
FvVQ21	*FaVQ21A*	*augustus_masked-Fvb6-2-processed-gene-63.18-mRNA-1*	97.11	0.0227	0.0478	0.4749
	FaVQ21B	*augustus_masked-Fvb6-3-processed-gene-154.9-mRNA-1*	98.13	0.0178	0.0229	0.7773
	FaVQ21C	*augustus_masked-Fvb6-4-processed-gene-221.8-mRNA-1*	97.62	0.023	0.0287	0.8014
	FaVQ21D	*augustus_masked-Fvb6-1-processed-gene-159.12-mRNA-1*	99.66	0.0047	0	NA
FvVQ22	*FaVQ22A*	*augustus_masked-Fvb6-2-processed-gene-358.15-mRNA-1*	96.29	0.0518	0.096	0.5396
	FaVQ22B	*augustus_masked-Fvb6-3-processed-gene-81.16-mRNA-1*	97.32	0.0386	0.1076	0.3587
	FaVQ22C	*augustus_masked-Fvb6-4-processed-gene-287.2-mRNA-1*	97.15	0.0356	0.1018	0.3497
	FaVQ22D	*augustus_masked-Fvb6-1-processed-gene-89.12-mRNA-1*	99.43	0.0252	0.061	0.4131
FvVQ23	*FaVQ23A*	*augustus_masked-Fvb6-2-processed-gene-357.6-mRNA-1*	95.74	0.0469	0.1798	0.2608
	FaVQ23B	*augustus_masked-Fvb6-3-processed-gene-80.10-mRNA-1*	92.88	0.0533	0.2799	0.1904
	FaVQ23D	*augustus_masked-Fvb6-1-processed-gene-88.4-mRNA-1*	96.64			
FvVQ24	*FaVQ24A*	*augustus_masked-Fvb6-2-processed-gene-353.4-mRNA-1*	97.38	0.0093	0.0548	0.1697
	FaVQ24B	*augustus_masked-Fvb6-3-processed-gene-76.4-mRNA-1*	97.25	0.0092	0.0707	0.1301
	FaVQ24C	*augustus_masked-Fvb6-4-processed-gene-291.7-mRNA-1*	96.88	0.0038	0.0567	0.0670
	FaVQ24D	*augustus_masked-Fvb6-1-processed-gene-81.7-mRNA-1*	99.86	0.0021	0	NA
FvVQ25	*FaVQ25A*	*snap_masked-Fvb7-1-processed-gene-32.10-mRNA-1*	96.21	0.0155	0.1095	0.1416
	FaVQ25B	*augustus_masked-Fvb7-3-processed-gene-193.5-mRNA-1*	95.27			
	FaVQ25C	*augustus_masked-Fvb7-4-processed-gene-191.1-mRNA-1*	95.77			
	FaVQ25D	*augustus_masked-Fvb7-2-processed-gene-32.8-mRNA-1*	100	0	0	NA

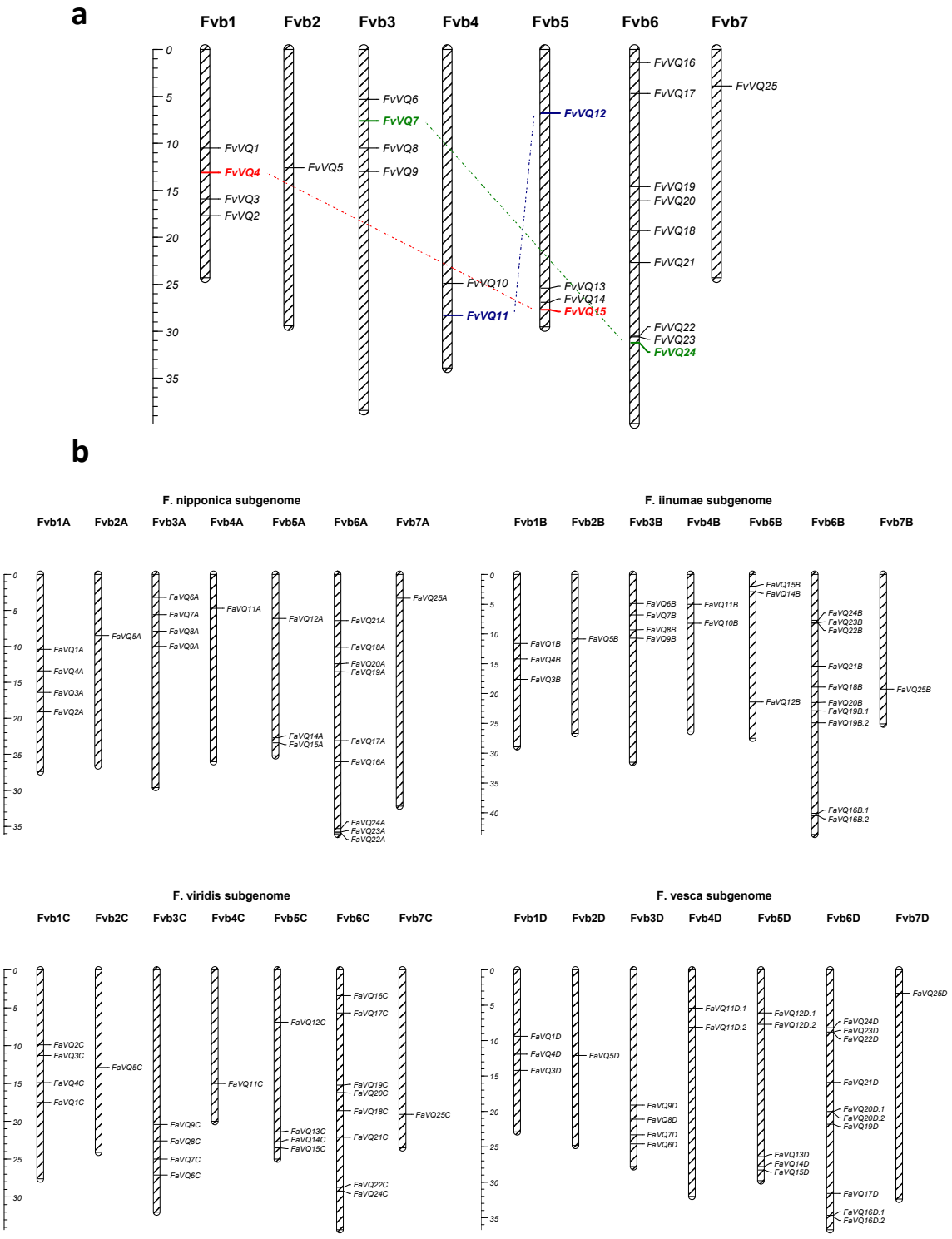


Figura 10. Mapa cromosómico de la familia VQ de *F. vesca* (a) y *F. x ananassa* (b). Las duplicaciones génicas que conservan la colinealidad se representan coloreadas y están conectadas mediante lineas. Las barras de escala representan el tamaño en Mb.

Algunas proteínas VQ de *F. x ananassa* poseen dominios adicionales al dominio VQ (Figura 11). Así, al igual que en FvVQ13, en las proteínas ortólogas FaVQ13C y FaVQ13D se identifican dominios NB-ARC y LRR_8 característicos de proteínas R. Además, dos proteínas VQ de *F. x ananassa* poseen dominios únicos adicionales, no encontrados en la familia VQ de *F. vesca*: FaVQ17C y FaVQ17D, que portan dominios Peptidase_M24 (Metallopeptidase family M24; PF00557) y AMP_N (Aminopeptidase P, N-terminal domain; PF05195). La proteína FaVQ17D incluye, además, un dominio ADK (Adenylate kinase; PF00406) y dos dominios DUF538 (PF04398) de función desconocida. Por último, FaVQ19A posee un dominio FAD-oxidase_C (FAD linked oxidases, C-terminal domain; PF02913). Ninguno de estos dominios es encontrado, sin embargo, en los ortólogos FvVQ17 o FvVQ19, ni en los homoeólogos correspondientes, lo que sugiere que la formación de estas proteínas quiméricas puede ser consecuencia de la acción discreta de elementos transponibles o de raros eventos de recombinación ectópica (Vicient and Casacuberta, 2017;Bailey et al., 2018). La búsqueda de proteínas portadoras de combinaciones de dominios similares, mediante CDART, no arrojó ningún resultado, por lo que, de confirmarse su expresión en futuros trabajos, se trataría de proteínas únicas en el reino vegetal.

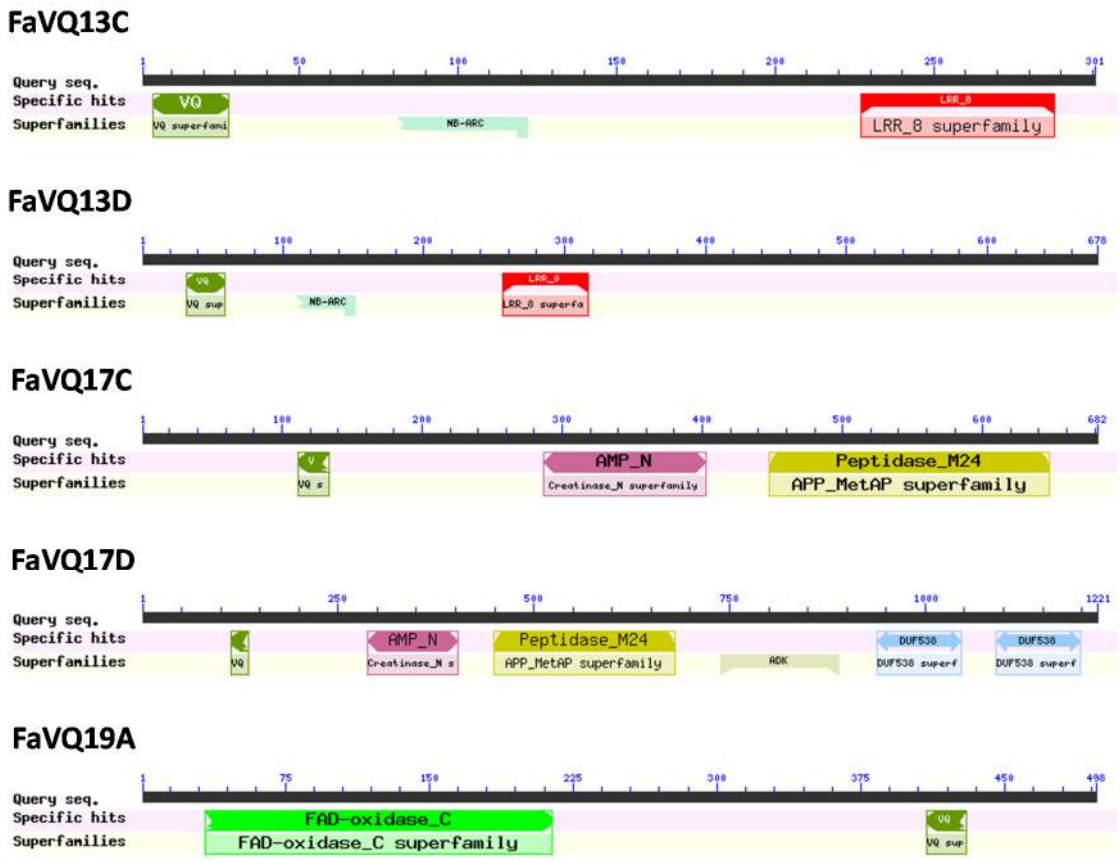


Figura 11. Dominios adicionales encontrados en algunas proteínas VQ de *F. x ananassa*.

3.6.2.2. Filogenómica y evolución de la familia VQ de fresa

El estudio del origen y evolución de la familia VQ de fresa se ha abordado mediante la búsqueda de ortólogos y análisis de sintenia y colinealidad con otras especies, aunque principalmente con la planta modelo *A. thaliana*.

Para la búsqueda y clasificación de las proteínas VQ ortólogas en la fresa diploide y octoploide, así como en otras especies, se usó la base de datos precomputados eggNOG, que en su última versión (EggNOG v5.0) incluye varios millones de proteínas, procedentes de más de 25.000 genomas distribuidos en 379 niveles taxonómicos, siendo por tanto una de las bases de datos de ortólogos más completas actualmente (Huerta-Cepas et al., 2019). La Tabla 7 recoge la clasificación de las familias FvVQ y FaVQ en diferentes grupos de ortólogos (*ortologous groups*, OGs), así como los ortólogos y un análisis de los mejores recíprocos mediante BLASTP encontrados en *A. thaliana*, que son los mejor caracterizados funcionalmente hasta la fecha.

Encontramos que las proteínas VQ ortólogas de *F. vesca* y *F. x ananassa* comparten la clasificación dentro de los mismos grupos de ortólogos, excepto las proteínas quiméricas FaVQ17C, FaVQ17D y FaVQ19A, exclusivas de *F. x ananassa*. La asignación de ortólogos en *A. thaliana*, considerando los OGs propios del grupo Viridiplantae, coincide esencialmente con la descrita previamente (Table 3) aunque con algunas diferencias. Así, las proteínas FvVQ5, FvVQ7, FvVQ9, FvVQ15, FvVQ16, FvVQ19 y FvVQ20, aparecen agrupados en OGs que incluyen proteínas AtVQ distintas a las anteriores. Aunque en el caso de FvVQ15 ello podría atribuirse a las diferencias encontradas en la secuencia proteica entre las dos versiones de anotación utilizadas (ver Tabla 5), este hecho se debe, más probablemente, a diferencias metodológicas.

Para despejar estas incógnitas, se realizó un estudio de las duplicaciones génicas de las familias VQ de *F. vesca* y *A. thaliana* , así como la sintenia conservada con los genes VQ de *A. thaliana* (Tabla 8), acompañado de un estudio filogenético de las proteínas VQ de *F. vesca*, *F. x ananassa* y *A. thaliana* (Figura 12) que nos permitiera establecer las relaciones evolutivas entre estas especies. Así, para FvVQ5 y FvVQ7, el estudio de colinealidad con *A. thaliana* no aporta ningún resultado, como tampoco el de reciprocidad por BLASTP. Sin embargo, en el caso de FvVQ5, la distribución filogenética es compatible con la atribución de AtVQ3 como ortólogo, además de que AtVQ8 y AtVQ20 resultan como posibles ortólogos de otras FvVQ. En el caso de FvVQ7, tanto nuestro análisis filogenético, como el resultado de eggNOG basado en un alineamiento filogenético más amplio (57 proteínas en 36 especies) coinciden en la asignación de AtVQ28 como el ortólogo más probable, contrastando así con los resultados previos. Por otra parte, para *FvVQ9* y *FvVQ20* encontramos relaciones colineales con *AtVQ19* y *AtVQ17*, respectivamente, coincidiendo además con la asignación de ortólogos de eggNOG. En el caso de FvVQ16, FvVQ19, FvVQ22 y FvVQ23, el análisis filogenómico no ofrece resultados definitivos sobre los posibles ortólogos en *Arabidopsis*.

Tabla 7. Clasificación, ortología y anotación de las proteínas VQ de fresa en eggNOG. Los mejores recíprocos mediante BLASTP están marcados por un asterisco.

Proteína	eggNOG OG por clade					Ortólogo en At	Anotación
	root	Eukaryota	Viridiplantae	Streptophyta	fabids		
Fv/FaVQ1	2E3PZ	2SAQX	37WJV	3GKNI	4JQXG	AtVQ12*	VQ motif
Fv/FaVQ2	2BI8E	2S1DP	37V7N	3GJ0U		AtVQ31*	VQ motif-containing protein
Fv/FaVQ3	2BZ6W	2S2J0	37V6K	3GJSP	4JPZA	AtVQ21* (MKS1)	VQ motif
Fv/FaVQ4	2C620	2QUSJ	37JW9	3GH2F	4JNWM	AtVQ34*	VQ motif
Fv/FaVQ5	2CA21	2S37N	37UZN	3GI6U	4JPHY	AtVQ3, AtVQ8, AtVQ20	VQ motif
Fv/FaVQ6	29BAJ	2RIDK	37RWG	3GESS	4JTXC	AtVQ4*, AtVQ13	VQ motif
Fv/FaVQ7	2CNAD	2S40B	37WM4	3GK5P	4JR2R	AtVQ28	VQ motif
Fv/FaVQ8	2C2QB	2SFRK	37XF4	3GKZR	4JR5T	AtVQ10*	VQ motif
Fv/FaVQ9	29BAJ	2RIDK	37RWG	3GGYR	4JNZQ	AtVQ19, AtVQ33	VQ motif
Fv/FaVQ10	2CMZQ	2QSYF	37SW9	3GHBQ	4JFSV	AtVQ6*, AtVQ14 (IKU1)	VQ motif
Fv/FaVQ11	2CA21	2S37N	37UZN	3GHXC	4JUII	AtVQ3, AtVQ8*, AtVQ20	VQ motif
Fv/FaVQ12	2CA21	2S37N	37UZN	3GHXC	4JPZ8	AtVQ3, AtVQ8, AtVQ20	VQ motif
Fv/FaVQ13	KOG4658	KOG4658	37YY9	3GP2S			ADP binding
Fv/FaVQ14	2CNBS	2QV2B	37SX1	3GCDN	4JE01	AtVQ9*	VQ motif
Fv/FaVQ15	2C620	2QUSJ	37JW9	3GH2F		AtVQ7*, AtVQ34	VQ motif-containing protein
Fv/FaVQ16	2E6SB	2SDF3	37XPP	3GMFT	4JR8U		VQ motif
Fv/FaVQ17	2C47I	2S2VA	37VUH	3GIEA		AtVQ24*	VQ motif
Fv/FaVQ18	2CJ1Z	2SFCE	37XQS	3GM0R		AtVQ16 (SIB2)	Sigma factor binding protein
Fv/FaVQ19	2E9ZH	2SG9H	37XN9	3GMA1	4JV57		VQ motif
Fv/FaVQ20	2CBNJ	2S3AS	37VSN	3GK2F	4JQUH	AtVQ17, AtVQ25	VQ motif
Fv/FaVQ21	2BWF4	2S350	37X84	3GJB0	4JQ2N	AtVQ11*	VQ motif
Fv/FaVQ22	2E3H3	2SAJD	37WYS	3GKX1	4JQY3		VQ motif
Fv/FaVQ23	2E3H3	2SAJD	37WYS	3GKX1	4JQY3		VQ motif
Fv/FaVQ24	2D220	2S4XE	37WDW	3GJQK		AtVQ22, AtVQ27	VQ motif-containing protein
Fv/FaVQ25	2CJ1Z	2S9W6	37WR8	3GMT7	4JQWF	AtVQ23* (SIB1)	Sigma factor binding protein
FaVQ17C	COG0006	KOG2414	37HHB	3GA57	4JEXN	AT1G09300	Xaa-Pro aminopeptidase 3
FaVQ17D	COG0006	KOG2414	37HHB	3GA57	4JEXN	AT1G09300	Xaa-Pro aminopeptidase 3
FaVQ19A	COG0277	KOG1231	37P3I	3G9BY	4JE9S	AT5G06580	D-lactate dehydrogenase cytochrome

Tabla 8. Duplicaciones génicas de los genes VQ de *F. vesca* y *A. thaliana* y análisis de sintenia entre las familias *FvVQ* y *AtVQ*. Estos análisis pueden replicarse siguiendo los enlaces persistentes indicados en la sección metodológica. El análisis adicional de auto-sintenia en *A. thaliana* puede replicarse en el siguiente enlace: https://genomevolution.org/r/171hr.

Gen 1	Gen2	Tipo	Kn	Ks	ω
FvVQ4	FvVQ15	Segmental	1.1604	7.0931	0.1636
FvVQ7	FvVQ24	Segmental	0.8222	60.3925	0.0136
FvVQ11	FvVQ12	Segmental	0.7355	8.5527	0.0860
AtVQ1	AtVQ10	Segmental	0.2183	0.6521	0.3348
AtVQ4	AtVQ13	Segmental	NA	NA	NA
AtVQ6	AtVQ14	Segmental	NA	NA	NA
AtVQ12	AtVQ29	Segmental	0.3212	1.2338	0.2603
AtVQ15	AtVQ24	Segmental	0.3005	1.3916	0.2159
AtVQ16	AtVQ23	Segmental	0.2953	0.9138	0.3232
AtVQ18	AtVQ26	Segmental	1.1247	67.6547	0.0166
AtVQ22	AtVQ27	Segmental	0.3015	1.1366	0.2653
AtVQ30	AtVQ34	Segmental	0.519	3.9624	0.1310
AtVQ2	AtVQ3	Tandem	NA	NA	NA
FvVQ4	AtVQ34	Colineal	0.9969	64.3349	0.0155
FvVQ4	AtVQ30	Colineal	1.1894	7.0478	0.1688
FvVQ8	AtVQ10	Colineal	1.3432	7.473	0.1797
FvVQ8	AtVQ1	Colineal	1.3791	3.4121	0.4042
FvVQ9	AtVQ19	Colineal	0.4417	5.6186	0.0786
FvVQ11	AtVQ8	Colineal	1.1552	68.5818	0.0168
FvVQ12	AtVQ20	Colineal	1.5475	63.4676	0.0244
FvVQ15	AtVQ7	Colineal	0.8568	6.1176	0.1401
FvVQ18	AtVQ16	Colineal	0.6835	68.1414	0.0100
FvVQ20	AtVQ17	Colineal	0.4996	2.5473	0.1961
FvVQ24	AtVQ27	Colineal	0.7464	59.4719	0.0126

Cabe mencionar el caso de *FvVQ22* y *FvVQ23* ya que, aunque no fueron detectados como duplicaciones génicas, existe una probabilidad elevada de que se trate de genes duplicados traspuestos, ya que ambos genes comparten una alta identidad (91,73%) y se encuentran en localizaciones cromosómicas relativamente próximas (unas 172 Kb) y orientados en sentido inverso. De esta forma, al menos uno de estos dos genes parece haber sido duplicado e insertado en otra localización cromosómica, posiblemente debido a la acción de elementos transponibles, rompiéndose así la colinealidad entre ellos. Los resultados previos indicaban que estos genes no eran expresados en los frutos ni tejidos vegetativos testados de *F. vesca* cv. Reina de los Valles y *F. x ananassa* cv. Camarosa. Los datos de expresión mediante análisis de RNA-seq, realizados en 46 tejidos de *F. vesca* y publicados recientemente (Li et al., 2019) revelan, sin embargo, que *FvVQ23* parece ser discretamente expresado en algunos tejidos (Figura 13).

En la Tabla 9 se sumariza la nueva asignación de probales ortólogos encontrados entre las familias VQ de *F. vesca* y *A. thaliana* según las evidencias descritas. En algunos casos no fue posible dicha asignación, mientras que en otros se señalan dos AtVQs como probables ortólogos. Este último es el caso de FvVQ6 y FvVQ10, cuyos posibles ortólogos resultan ser duplicaciones génicas en Arabidopsis (ver Tabla 8).

Tabla 9. Nueva asignación de ortólogos entre las familias VQ de *F. vesca* y *A. thaliana*. Se indica el tipo de evidencia en que se sustenta la asignación (p: homología a nivel de proteína; s: sintenia conservada). Las proteínas detectadas como mejores recíprocos mediante BLASTP se marcan con un asterisco.

FvVQ	Ortólogo en Arabidopsis	Evidencia
FvVQ1	AtVQ12*	p
FvVQ2	AtVQ31*	p
FvVQ3	AtVQ21* (MKS1)	p
FvVQ4	AtVQ34*	p,s
FvVQ5	AtVQ3	p
FvVQ6	AtVQ4*, AtVQ13	p
FvVQ7	AtVQ28	p
FvVQ8	AtVQ1, AtVQ10*	p,s
FvVQ9	AtVQ19	p,s
FvVQ10	AtVQ6*, AtVQ14 (IKU1)	p
FvVQ11	AtVQ8*	p,s
FvVQ12	AtVQ20	p,s
FvVQ13		
FvVQ14	AtVQ9*	p
FvVQ15	AtVQ7*	p,s
FvVQ16		
FvVQ17	AtVQ24*	p
FvVQ18	AtVQ16 (SIB2)	p,s
FvVQ19		
FvVQ20	AtVQ17	p,s
FvVQ21	AtVQ11*	p
FvVQ22		
FvVQ23		
FvVQ24	AtVQ22	p,s
FvVQ25	AtVQ23* (SIB1)	p

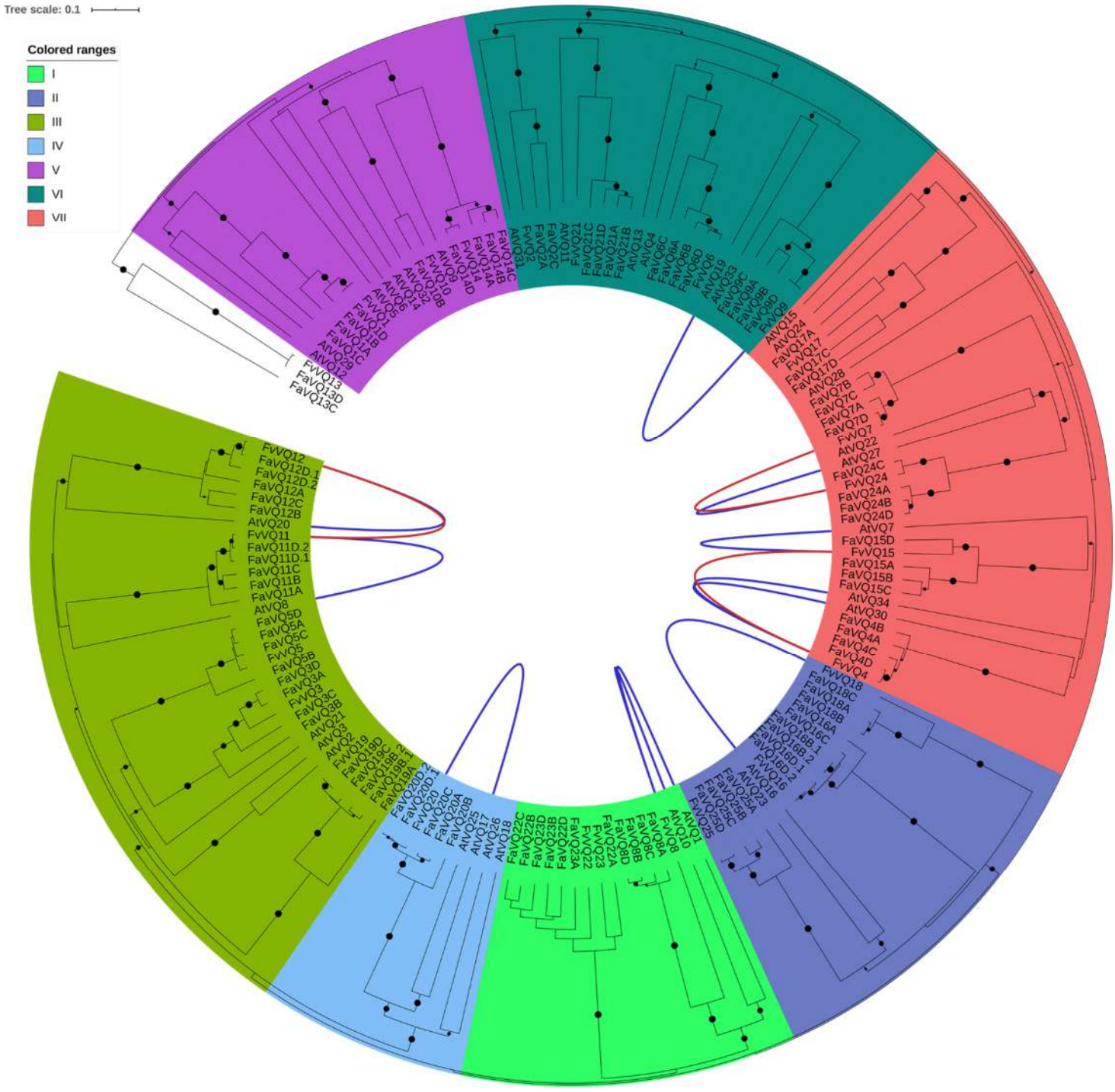


Figura 12. Árbol filogenético (N-J, 1000 replicados) de las proteínas VQ de *F. vesca*, *F. x ananassa* y *A. thaliana*. La escala evolutiva, computada mediante el método *p-distance*, se mide en substituciones de aminoácidos por sitio. Las relaciones sinténicas entre distintos genes están representadas por líneas (rojo en el caso de parálogos de *F. vesca* y azul para ortólogos colineales entre *F. vesca* y *A. thaliana*). Se han mantenido los colores asignados a cada grupo previamente como referencia. El área de los círculos en cada rama simboliza el soporte obtenido mediante *bootstrapping* (50-100%).

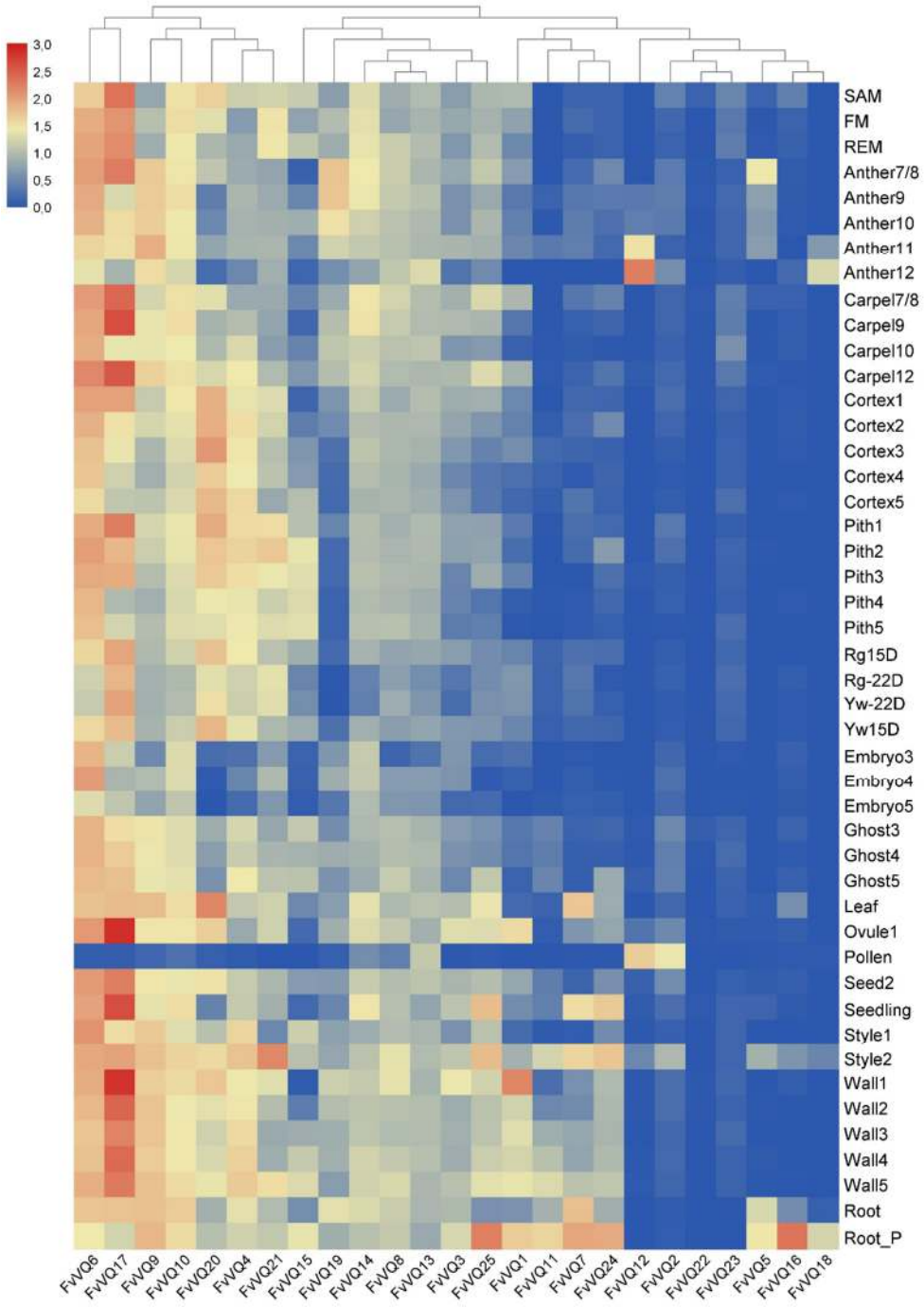


Figura 13. Expresión de los genes *FvVQ* en diferentes tejidos de *F. vesca*. Los datos de expresión (TPM) fueron tomados de Li et al., 2019 (Tabla suplementaria S6). La escala de color muestra los valores logarítmicos (log_{10}) de expresión. La descripción de los tejidos representados puede encontrarse en la Figura 15 del Capítulo 2.

3.6.2.3. Interacciones proteína-proteína entre las familias VQ y WRKY de fresa en respuesta a *C. acutatum*.

Debido a las nuevas asignaciones obtenidas de probables ortólogos entre las familias de VQ y WRKY de fresa y *Arabidopsis*, es necesario introducir ciertas modificaciones en la anterior Figura 8. Además, hemos incluido al ortólogo en fresa de *WRKY75*, *Fv/FaWRKY24* (denominado en trabajos previos *FaWRKY1*) (Encinas-Villarejo et al., 2009;Amil-Ruiz et al., 2016;Higuera et al., 2019) dado que también juega un papel en la defensa frente a *C. acutatum*, a la vez que forma parte de la intrincada red de interacciones proteína-proteína de la familia VQ. Los resultados se representan en la Figura 16, donde encontramos que los *clusters* de interacciones centradas en los ortólogos de WRKY25-WRKY33 y WRKY75 se relacionan a través de SIB1 y SIB2, lo cual podría crear una competencia entre los factores WRKY33 y WRKY75 por la unión con SIB1 y SIB2 en aquellas condiciones de estrés reguladas por dichos factores, como por ejemplo, la respuesta de defensa frente a *C. acutatum*, que merecería la pena explorar en futuros trabajos.

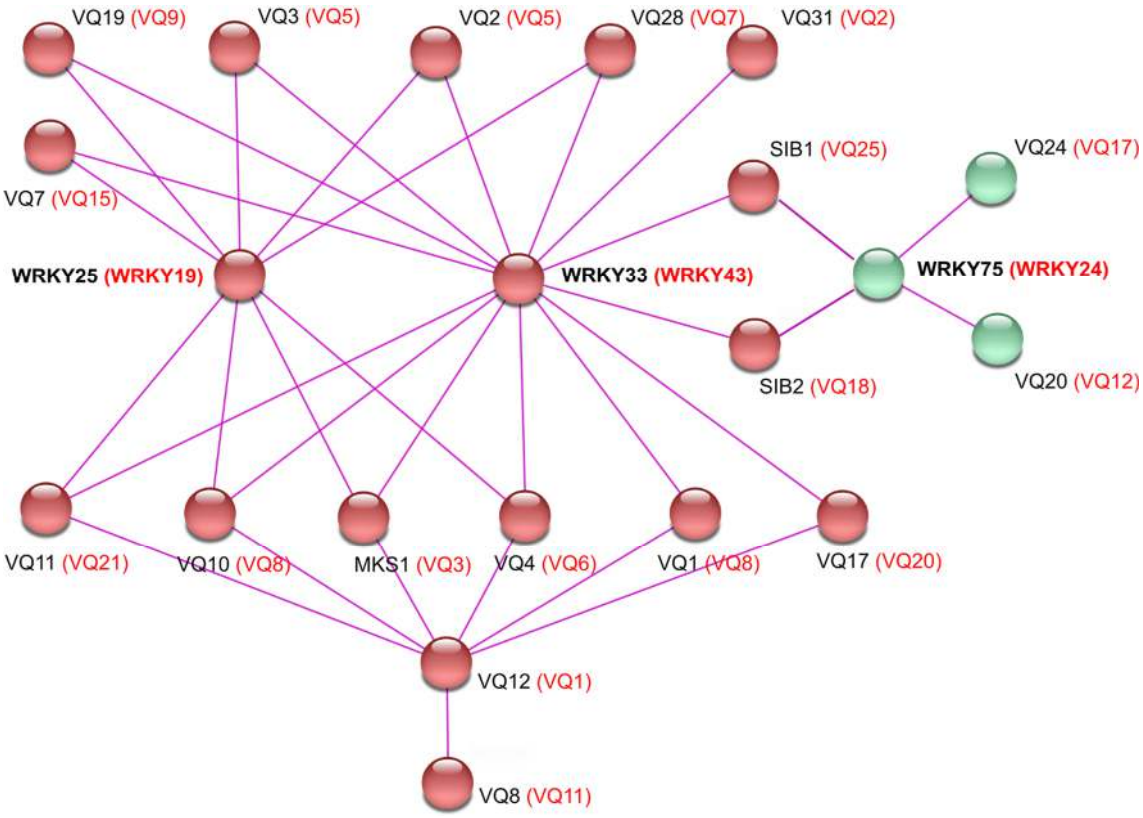


Figura 16. Redes de interacciones proteína-proteína entre la familia VQ y los factores de transcripción WRKY25, WRKY33 y WRKY75. El esquema fue construido en STRING versión 11.0 y corregido en base a las interacciones conocidas entre proteínas VQ y factores WRKY de *A. thaliana* citados en el texto (anotados en color negro), con sus correspondientes ortólogos en fresa (entre paréntesis y anotados en color rojo). Los grupos de interacción VQ-WRKY25-33 y VQ-WRKY75 se representan en color rojo y verde, respectivamente.

3.7. Bibliografía

Alves, M.S., Dadalto, S.P., Goncalves, A.B., De Souza, G.B., Barros, V.A., and Fietto, L.G. (2014). Transcription Factor Functional Protein-Protein Interactions in Plant Defense Responses. *Proteomes* 2**,** 85-106.doi: 10.3390/proteomes2010085

Amil-Ruiz, F., Garrido-Gala, J., Blanco-Portales, R., Folta, K.M., Munoz-Blanco, J., and Caballero, J.L. (2013). Identification and validation of reference genes for transcript normalization in strawberry (Fragaria x ananassa) defense responses. *PLoS One* 8**,** e70603.doi: 10.1371/journal.pone.0070603

Amil-Ruiz, F., Garrido-Gala, J., Gadea, J., Blanco-Portales, R., Munoz-Merida, A., Trelles, O., et al. (2016). Partial Activation of SA- and JA-Defensive Pathways in Strawberry upon Colletotrichum acutatum Interaction. *Front Plant Sci* 7**,** 1036.doi: 10.3389/fpls.2016.01036

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A., et al. (2016). Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* 44**,** W147-153.doi: 10.1093/nar/gkw419

Bailey, P.C., Schudoma, C., Jackson, W., Baggs, E., Dagdas, G., Haerty, W., et al. (2018). Dominant integration locus drives continuous diversification of plant immune receptors with exogenous domain fusions. *Genome Biol* 19**,** 23.doi: 10.1186/s13059-018-1392-6

Bakshi, M., and Oelmuller, R. (2014). WRKY transcription factors: Jack of many trades in plants. *Plant Signal Behav* 9**,** e27700.doi: 10.4161/psb.27700

Birkenbihl, R.P., Diezel, C., and Somssich, I.E. (2012). Arabidopsis WRKY33 is a key transcriptional regulator of hormonal and metabolic responses toward Botrytis cinerea infection. *Plant Physiol* 159**,** 266-285.doi: 10.1104/pp.111.192641

Chen, J., Wang, H., Li, Y., Pan, J., Hu, Y., and Yu, D. (2018). Arabidopsis VQ10 interacts with WRKY8 to modulate basal defense against Botrytis cinerea. *J Integr Plant Biol* 60**,** 956-969.doi: 10.1111/jipb.12664

Cheng, Y., Zhou, Y., Yang, Y., Chi, Y.J., Zhou, J., Chen, J.Y., et al. (2012). Structural and functional analysis of VQ motif-containing proteins in Arabidopsis as interacting proteins of WRKY transcription factors. *Plant Physiol* 159**,** 810-825.doi: 10.1104/pp.112.196816

Chu, W., Liu, B., Wang, Y., Pan, F., Chen, Z., Yan, H., et al. (2016). Genome-wide analysis of poplar VQ gene family and expression profiling under PEG, NaCl, and SA treatments. *Tree Genetics & Genomes* 12.doi: 10.1007/s11295-016-1082-z

Csukasi, F., Osorio, S., Gutierrez, J.R., Kitamura, J., Giavalisco, P., Nakajima, M., et al. (2011). Gibberellin biosynthesis and signalling during development of the strawberry receptacle. *New Phytol* 191**,** 376-390.doi: 10.1111/j.1469-8137.2011.03700.x

Darwish, O., Shahan, R., Liu, Z., Slovin, J.P., and Alkharouf, N.W. (2015). Re-annotation of the woodland strawberry (Fragaria vesca) genome. *BMC Genomics* 16**,** 29.doi: 10.1186/s12864-015-1221-1

Denoyes-Rothan, B., Guerin, G., Delye, C., Smith, B., Minz, D., Maymon, M., et al. (2003). Genetic diversity and pathogenic variability among isolates of colletotrichum species from strawberry. *Phytopathology* 93**,** 219-228.doi: 10.1094/PHYTO.2003.93.2.219

Dimeglio, L.M., Staudt, G., Yu, H., and Davis, T.M. (2014). A phylogenetic analysis of the genus Fragaria (strawberry) using intron-containing sequence from the ADH-1 gene. *PLoS One* 9**,** e102237.doi: 10.1371/journal.pone.0102237

Dong, Q., Zhao, S., Duan, D., Tian, Y., Wang, Y., Mao, K., et al. (2018). Structural and functional analyses of genes encoding VQ proteins in apple. *Plant Sci* 272**,** 208-219.doi: 10.1016/j.plantsci.2018.04.029

Edger, P.P., Poorten, T.J., Vanburen, R., Hardigan, M.A., Colle, M., Mckain, M.R., et al. (2019). Origin and evolution of the octoploid strawberry genome. *Nat Genet* 51**,** 541-547.doi: 10.1038/s41588-019-0356-4

Edger, P.P., Vanburen, R., Colle, M., Poorten, T.J., Wai, C.M., Niederhuth, C.E., et al. (2018). Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (Fragaria vesca) with chromosome-scale contiguity. *Gigascience* 7**,** 1-7.doi: 10.1093/gigascience/gix124

Encinas-Villarejo, S., Maldonado, A.M., Amil-Ruiz, F., De Los Santos, B., Romero, F., Pliego-Alfaro, F., et al. (2009). Evidence for a positive regulatory role of strawberry (Fragaria×ananassa) Fa WRKY1 and Arabidopsis At WRKY75 proteins in resistance. *Journal of Experimental Botany* 60**,** 3043-3065.doi: 10.1093/jxb/erp152

Eulgem, T., Rushton, P.J., Robatzek, S., and Somssich, I.E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci* 5**,** 199-206.doi: 10.1016/s1360-1385(00)01600-9

Garner, C.M., Kim, S.H., Spears, B.J., and Gassmann, W. (2016). Express yourself: Transcriptional regulation of plant innate immunity. *Semin Cell Dev Biol* 56**,** 150-162.doi: 10.1016/j.semcdb.2016.05.002

Guidarelli, M., Carbone, F., Mourgues, F., Perrotta, G., Rosati, C., Bertolini, P., et al. (2011). Colletotrichum acutatum interactions with unripe and ripe strawberry fruits and differential responses at histological and transcriptional levels. *Plant Pathology* 60**,** 685-697.doi: 10.1111/j.1365-3059.2010.02423.x

Guo, J., Chen, J., Yang, J., Yu, Y., Yang, Y., and Wang, W. (2018). Identification, characterization and expression analysis of the VQ motif-containing gene family in tea plant (Camellia sinensis). *BMC Genomics* 19**,** 710.doi: 10.1186/s12864-018-5107-x

Haug-Baltzell, A., Stephens, S.A., Davey, S., Scheidegger, C.E., and Lyons, E. (2017). SynMap2 and SynMap3D: web-based whole-genome synteny browsers. *Bioinformatics* 33**,** 2197-2198.doi: 10.1093/bioinformatics/btx144

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., and Vandesompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* 8**,** R19.doi: 10.1186/gb-2007-8-2-r19

Higuera, J.J., Garrido-Gala, J., Lekhbou, A., Arjona-Girona, I., Amil-Ruiz, F., Mercado, J.A., et al. (2019). The Strawberry FaWRKY1 Transcription Factor Negatively Regulates Resistance to Colletotrichum acutatum in Fruit Upon Infection. *Front Plant Sci* 10**,** 480.doi: 10.3389/fpls.2019.00480

Hirakawa, H., Shirasawa, K., Kosugi, S., Tashiro, K., Nakayama, S., Yamada, M., et al. (2014). Dissection of the octoploid strawberry genome by deep sequencing of the genomes of Fragaria species. *DNA Res* 21**,** 169-181.doi: 10.1093/dnares/dst049

Hu, B., Jin, J., Guo, A.Y., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31**,** 1296-1297.doi: 10.1093/bioinformatics/btu817

Hu, P., Zhou, W., Cheng, Z., Fan, M., Wang, L., and Xie, D. (2013). JAV1 controls jasmonate-regulated plant defense. *Mol Cell* 50**,** 504-515.doi: 10.1016/j.molcel.2013.04.027

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., et al. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* 44**,** D286-293.doi: 10.1093/nar/gkv1248

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernandez-Plaza, A., Forslund, S.K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 47**,** D309-D314.doi: 10.1093/nar/gky1085

Hulsen, T., De Vlieg, J., and Alkema, W. (2008). BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* 9**,** 488.doi: 10.1186/1471-2164-9-488

Jiang, J., Ma, S., Ye, N., Jiang, M., Cao, J., and Zhang, J. (2017). WRKY transcription factors in plant responses to stresses. *J Integr Plant Biol* 59**,** 86-101.doi: 10.1111/jipb.12513

Jiang, S.Y., Sevugan, M., and Ramachandran, S. (2018). Valine-glutamine (VQ) motif coding genes are ancient and non-plant-specific with comprehensive expression regulation by various biotic and abiotic stresses. *BMC Genomics* 19**,** 342.doi: 10.1186/s12864-018-4733-7

Jing, Y., and Lin, R. (2015). The VQ Motif-Containing Protein Family of Plant-Specific Transcriptional Regulators. *Plant Physiol* 169**,** 371-378.doi: 10.1104/pp.15.00788

Jung, S., Lee, T., Cheng, C.H., Buble, K., Zheng, P., Yu, J., et al. (2019). 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. *Nucleic Acids Res* 47**,** D1137-D1145.doi: 10.1093/nar/gky1000

Kallberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., et al. (2012). Template-based protein structure modeling using the RaptorX web server. *Nat Protoc* 7**,** 1511-1522.doi: 10.1038/nprot.2012.085

Kim, D.Y., Kwon, S.I., Choi, C., Lee, H., Ahn, I., Park, S.R., et al. (2013). Expression analysis of rice VQ genes in response to biotic and abiotic stresses. *Gene* 529**,** 208-214.doi: 10.1016/j.gene.2013.08.023

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33**,** 1870-1874.doi: 10.1093/molbev/msw054

Lai, Z., Li, Y., Wang, F., Cheng, Y., Fan, B., Yu, J.Q., et al. (2011). Arabidopsis sigma factor binding proteins are activators of the WRKY33 transcription factor in plant defense. *Plant Cell* 23**,** 3824-3841.doi: 10.1105/tpc.111.090571

Le Berre, J.Y., Gourgues, M., Samans, B., Keller, H., Panabieres, F., and Attard, A. (2017). Transcriptome dynamic of Arabidopsis roots infected with Phytophthora parasitica identifies VQ29, a gene induced during the penetration and involved in the restriction of infection. *PLoS One* 12**,** e0190341.doi: 10.1371/journal.pone.0190341

Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47**,** W256-W259.doi: 10.1093/nar/gkz239

Li, N., Li, X., Xiao, J., and Wang, S. (2014a). Comprehensive analysis of VQ motif-containing gene expression in rice defense responses to three pathogens. *Plant Cell Rep* 33**,** 1493-1505.doi: 10.1007/s00299-014-1633-4

Li, Y., Jing, Y., Li, J., Xu, G., and Lin, R. (2014b). Arabidopsis VQ MOTIF-CONTAINING PROTEIN29 represses seedling deetiolation by interacting with PHYTOCHROME-INTERACTING FACTOR1. *Plant Physiol* 164**,** 2068-2080.doi: 10.1104/pp.113.234492

Li, Y., Pi, M., Gao, Q., Liu, Z., and Kang, C. (2019). Updated annotation of the wild strawberry Fragaria vesca V4 genome. *Hortic Res* 6**,** 61.doi: 10.1038/s41438-019-0142-6

Li, Y., Wei, W., Feng, J., Luo, H., Pi, M., Liu, Z., et al. (2018). Genome re-annotation of the wild strawberry Fragaria vesca using extensive Illumina- and SMRT-based RNA-seq datasets. *DNA Res* 25**,** 61-70.doi: 10.1093/dnares/dsx038

Lin, J.R., and Hu, J. (2013). SeqNLS: nuclear localization signal prediction based on frequent pattern mining and linear motif scoring. *PLoS One* 8**,** e76864.doi: 10.1371/journal.pone.0076864

Luhua, S., Ciftci-Yilmaz, S., Harper, J., Cushman, J., and Mittler, R. (2008). Enhanced tolerance to oxidative stress in transgenic Arabidopsis plants expressing proteins of unknown function. *Plant Physiol* 148**,** 280-292.doi: 10.1104/pp.108.124875

Lyons, E., and Freeling, M. (2008). How to usefully compare homologous plant genes and chromosomes as DNA sequences. *Plant J* 53**,** 661-673.doi: 10.1111/j.1365-313X.2007.03326.x

Machens, F., Becker, M., Umrath, F., and Hehl, R. (2014). Identification of a novel type of WRKY transcription factor binding site in elicitor-responsive cis-sequences from Arabidopsis thaliana. *Plant Mol Biol* 84**,** 371-385.doi: 10.1007/s11103-013-0136-y

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., et al. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45**,** D200-D203.doi: 10.1093/nar/gkw1129

Margara, J. (1989). *Bases de la multiplication végétative : les méristèmes et l'organogenèse.* Paris : Institut national de la recherche agronomique.

Okonechnikov, K., Golosova, O., Fursov, M., and Team, U. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 28**,** 1166-1167.doi: 10.1093/bioinformatics/bts091

Pazos, F., Pietrosemoli, N., Garcia-Martin, J.A., and Solano, R. (2013). Protein intrinsic disorder in plants. *Front Plant Sci* 4**,** 363.doi: 10.3389/fpls.2013.00363

Pecher, P., Eschen-Lippold, L., Herklotz, S., Kuhle, K., Naumann, K., Bethke, G., et al. (2014). The Arabidopsis thaliana mitogen-activated protein kinases MPK3 and MPK6 target a subclass of 'VQ-motif'-containing proteins to regulate immune responses. *New Phytol* 203**,** 592-606.doi: 10.1111/nph.12817

Pei, J., Kim, B.H., and Grishin, N.V. (2008). PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res* 36**,** 2295-2300.doi: 10.1093/nar/gkn072

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research* 29**,** e45-e45

Qiu, J.L., Fiil, B.K., Petersen, K., Nielsen, H.B., Botanga, C.J., Thorgrimsen, S., et al. (2008). Arabidopsis MAP kinase 4 regulates gene expression through transcription factor release in the nucleus. *EMBO J* 27**,** 2214-2221.doi: 10.1038/emboj.2008.147

Rinerson, C.I., Rabara, R.C., Tripathi, P., Shen, Q.J., and Rushton, P.J. (2015). The evolution of WRKY transcription factors. *BMC Plant Biol* 15**,** 66.doi: 10.1186/s12870-015-0456-y

Rousseau-Gueutin, M., Gaston, A., Ainouche, A., Ainouche, M.L., Olbricht, K., Staudt, G., et al. (2009). Tracking the evolutionary history of polyploidy in Fragaria L. (strawberry): new insights from phylogenetic analyses of low-copy nuclear genes. *Mol Phylogenet Evol* 51**,** 515-530.doi: 10.1016/j.ympev.2008.12.024

Rushton, P.J., Somssich, I.E., Ringler, P., and Shen, Q.J. (2010). WRKY transcription factors. *Trends Plant Sci* 15**,** 247-258.doi: 10.1016/j.tplants.2010.02.006

Schwab, W., Schaart, J.G., and Rosati, C. (2009). "Functional Molecular Biology Research in Fragaria," in *Genetics and Genomics of Rosaceae,* eds. K.M. Folta & S.E. Gardiner. (New York, NY: Springer New York), 457-486.

Shulaev, V., Sargent, D.J., Crowhurst, R.N., Mockler, T.C., Folkerts, O., Delcher, A.L., et al. (2011). The genome of woodland strawberry (Fragaria vesca). *Nat Genet* 43**,** 109-116.doi: 10.1038/ng.740

Sperschneider, J., Catanzariti, A.M., Deboer, K., Petre, B., Gardiner, D.M., Singh, K.B., et al. (2017). LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Sci Rep* 7**,** 44598.doi: 10.1038/srep44598

Symons, G.M., Chua, Y.J., Ross, J.J., Quittenden, L.J., Davies, N.W., and Reid, J.B. (2012). Hormonal changes during non-climacteric ripening in strawberry. *J Exp Bot* 63**,** 4741-4750.doi: 10.1093/jxb/ers147

Tennessen, J.A., Govindarajulu, R., Ashman, T.L., and Liston, A. (2014). Evolutionary origins and dynamics of octoploid strawberry subgenomes revealed by dense targeted capture linkage maps. *Genome Biol Evol* 6**,** 3295-3313.doi: 10.1093/gbe/evu261

Tsuda, K., and Somssich, I.E. (2015). Transcriptional networks in plant immunity. *New Phytol* 206**,** 932-947.doi: 10.1111/nph.13286

Van Verk, M.C., Pappaioannou, D., Neeleman, L., Bol, J.F., and Linthorst, H.J. (2008). A Novel WRKY transcription factor is required for induction of PR-1a gene expression by

salicylic acid and bacterial elicitors. *Plant Physiol* 146**,** 1983-1995.doi: 10.1104/pp.107.112789

Vicient, C.M., and Casacuberta, J.M. (2017). Impact of transposable elements on polyploid plant genomes. *Ann Bot* 120**,** 195-207.doi: 10.1093/aob/mcx078

Voorrips, R.E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 93**,** 77-78.doi: 10.1093/jhered/93.1.77

Wang, F., Zhang, F., Chen, M., Liu, Z., Zhang, Z., Fu, J., et al. (2017a). Comparative Transcriptomics Reveals Differential Gene Expression Related to Colletotrichum gloeosporioides Resistance in the Octoploid Strawberry. *Front Plant Sci* 8**,** 779.doi: 10.3389/fpls.2017.00779

Wang, H., Hu, Y., Pan, J., and Yu, D. (2015a). Arabidopsis VQ motif-containing proteins VQ12 and VQ29 negatively modulate basal defense against Botrytis cinerea. *Sci Rep* 5**,** 14185.doi: 10.1038/srep14185

Wang, M., Vannozzi, A., Wang, G., Zhong, Y., Corso, M., Cavallini, E., et al. (2015b). A comprehensive survey of the grapevine VQ gene family and its transcriptional correlation with WRKY proteins. *Front Plant Sci* 6**,** 417.doi: 10.3389/fpls.2015.00417

Wang, Y., Liu, H., Zhu, D., Gao, Y., Yan, H., and Xiang, Y. (2017b). Genome-wide analysis of VQ motif-containing proteins in Moso bamboo (Phyllostachys edulis). *Planta* 246**,** 165-181.doi: 10.1007/s00425-017-2693-9

Watson, R., Wright, C.J., Mcburney, T., Taylor, A.J., and Linforth, R.S. (2002). Influence of harvest date and light integral on the development of strawberry flavour compounds. *J Exp Bot* 53**,** 2121-2129.doi: 10.1093/jxb/erf088

Weyhe, M., Eschen-Lippold, L., Pecher, P., Scheel, D., and Lee, J. (2014). Menage a trois: the complex relationships between mitogen-activated protein kinases, WRKY transcription factors, and VQ-motif-containing proteins. *Plant Signal Behav* 9**,** e29519.doi: 10.4161/psb.29519

Wright, P.E., and Dyson, H.J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* 16**,** 18-29.doi: 10.1038/nrm3920

Yan, C., Fan, M., Yang, M., Zhao, J., Zhang, W., Su, Y., et al. (2018). Injury Activates Ca(2+)/Calmodulin-Dependent Phosphorylation of JAV1-JAZ8-WRKY51 Complex for Jasmonate Biosynthesis. *Mol Cell* 70**,** 136-149 e137.doi: 10.1016/j.molcel.2018.03.013

Yang, Z., and Nielsen, R. (2002). Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol* 19**,** 908-917.doi: 10.1093/oxfordjournals.molbev.a004148

Zhang, G., Wang, F., Li, J., Ding, Q., Zhang, Y., Li, H., et al. (2015). Genome-Wide Identification and Analysis of the VQ Motif-Containing Protein Family in Chinese Cabbage (Brassica rapa L. ssp. Pekinensis). *Int J Mol Sci* 16**,** 28683-28704.doi: 10.3390/ijms161226127

Zheng, Z., Qamar, S.A., Chen, Z., and Mengiste, T. (2006). Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant J* 48**,** 592-605.doi: 10.1111/j.1365-313X.2006.02901.x

Zhong, Y., Guo, C., Chu, J., Liu, H., and Cheng, Z.M. (2018). Microevolution of the VQ gene family in six species of Fragaria. *Genome* 61**,** 49-57.doi: 10.1139/gen-2017-0038

Zhou, Y., Yang, Y., Zhou, X., Chi, Y., Fan, B., and Chen, Z. (2016). Structural and Functional Characterization of the VQ Protein Family and VQ Protein Variants from Soybean. *Sci Rep* 6**,** 34663.doi: 10.1038/srep34663

Capítulo 4.

The strawberry FAWRKY1 transcription factor negatively regulates resistance to *Colletotrichum acutatum* in fruit upon infection

Este capítulo ha sido publicado el 18 de abril de 2019:

Higuera JJ*, Garrido-Gala J*, Lekhbou A, Arjona-Girona I, Amil-Ruiz F, Mercado JA, Pliego-Alfaro F, Muñoz-Blanco J, López-Herrera CJ and Caballero JL (2019). The Strawberry FaWRKY1 Transcription Factor Negatively Regulates Resistance to *Colletotrichum acutatum* in Fruit upon Infection. Front. Plant Sci. 10:480. doi: 10.3389/fpls.2019.00480

(*Ambos autores contribuyeron por igual a la realización de este trabajo)

Capítulo 4. The strawberry FAWRKY1 transcription factor negatively regulates resistance to *Colletotrichum acutatum* in fruit upon infection

4.1. Abstract

Strawberry (*Fragaria x ananassa*) is a major food crop worldwide, due to the flavor, aroma and health benefits of the fruit, but its productivity and quality are seriously limited by a large variety of phytopathogens, including *Colletotrichum* spp. So far, key factors regulating strawberry immune response remain unknown. The *FaWRKY1* gene has been previously proposed as an important element mediating defense responses in strawberry to *C. acutatum*. To get further insight into the functional role that *FaWRKY1* plays in the defense mechanism, *Agrobacterium*-mediated transient transformation was used both to silence and overexpress the *FaWRKY1* gene in strawberry fruits (*Fragaria x ananassa* cv. Primoris), which were later analyzed upon Colletotrichum acutatum inoculation. Susceptibility tests were performed after pathogen infection comparing the severity of disease between the two agroinfiltrated opposite halves of the same fruit, one half bearing a construct either for *FaWRKY1* overexpression or RNAi-mediated silencing and the other half bearing the empty vector, as control. The severity of tissue damage was monitored and found to be visibly reduced at five days after pathogen inoculation in the fruit half where *FaWRKY1* was transiently silenced compared to that of the opposite control half and statistical analysis corroborated a significant reduction in disease susceptibility. Contrarily, a similar level of susceptibility was found when *FaWRKY1* overexpression and control fruit samples, was compared. These results unravel a negative regulatory role of FaWRKY1 in resistance to the phytopathogenic fungus *Colletotrichum acutatum* in strawberry fruit and contrast with the previous role described for this gene in *Arabidopsis* as positive regulator of resistance against the bacteria *Pseudomonas syringae*. Based on previous results, a tentative working model for *WRKY75-like* genes after pathogen infection is proposed and the expression pattern of potential downstream FaWRKY1 target genes was also analyzed in strawberry fruit upon *C. acutatum* infection. Our results highlight that FaWRKY1 might display different function according to species, plant tissue and/or type of pathogen and underline the intricate FaWRKY1 responsive defense regulatory mechanism taking place in strawberry against this important crop pathogen.

4.2. Introduction

Strawberry has grown in importance throughout the world, due to both the fact that this small fruit has become a highly relevant product at the social level for its

nutritional properties and beneficial effects on health (Giampieri et al., 2017;Garrido-Bigotes et al., 2018), and its economic importance, being one of the products with the largest share in the export of fruits and vegetables (http://faostat.fao.org).

Strawberry exhibits wide diversity in its susceptibility to a large variety of phytopathogenic organisms, including *Colletotrichum* spp. causal agent of anthracnose, a major disease of this crop (Jeger and Bailey, 1992;Maas, 1998;Simpson, 2015), yielding major losses in fruit production at the pre-harvest stage (Guidarelli et al., 2011). Three Colletotrichum species cause the anthracnose diseases of strawberry: *C. fragariae* and *C. gloeosporioides* induce the crown rot and lesions in vegetative tissues, while *C. acutatum* is the main pathogen causing the strawberry fruit rot (Peres et al., 2005). Anthracnose on strawberry is found worldwide and is a very destructive disease, causing up to 80% of plant death in nurseries and over 50% of yield losses in fields (Sreenivasaprasad and Talhinhas, 2005). Fungicide applications are resource-demanding every year for this and other important crops, to secure production yield but this increases public concern for environment food safety and makes urgent the need to develop sustainable alternatives.

Colletotrichum spp. is considered a hemibiotrophic pathogen and the histopathology of the interaction strawberry-*C. acutatum* has been previously well reported and monitored by using light, fluorescent and transmission electron microscopy with a symptomless, brief biotrophic phase, preceding a main necrotrophic development within the strawberry tissues and the rise of lesions (Curry et al., 2002;Horowitz et al., 2002;Peres et al., 2005;Amil-Ruiz et al., 2016). Although much research has been dedicated to understand the interplay between fungal pathogen and plant, there is a lack of comprehensive information on the molecular level and mechanisms underlying the process of defense and resistance to this pathogen in strawberry. Thus, characterizing the strawberry defense-responsive components will lead to improve the understanding of the underlying molecular mechanisms of defense. Indeed, it represents a major piece within the strategies to improve resistance in this important crop, which is a relevant economic and environmental issue.

The resistance to invaders is frequently harmonized in plant through a complex defense molecular network fine-tuned by phytohormones such as salicylic acid (SA), jasmonic acid (JA) and ethylene (ET), mainly, but also abscisic acid (ABA) and gibberellins (GA), which regulate the defensive response to efficiently face the different pathogens (Robert-Seilaniantz et al., 2011). It is well known that SA signaling pathway is mainly activated against biotrophic pathogens, and often induces a hypersensitive response (HR) followed by the onset of Systemic Acquired Resistance (SAR) (Fu and Dong, 2013). On the contrary, JA/ET signaling pathway is commonly activated in plant against necrotrophic pathogens, insect, or in response to wounding (Yan and Xie, 2015). JA induces a different set of defense response genes and the production of a large variety of secondary metabolites such as alkaloids, phenolic compounds and terpenes (Zhou and Memelink, 2016). Crosstalk among these signaling

pathways has been well described in models (Robert-Seilaniantz et al., 2011) while remains largely unknown or poorly understood in many crop species.

Transcription factors (TFs) are key regulators of gene expression, which play important roles within this complex defense molecular network leading to plant immunity. To date, many defense-related TFs have been identified in plants, including MYBs, the TGA/bZIP family protein, AP2/ERF-ethylene responsive element binding factors, NACs, the Whirly (WHY) family protein, and WRKYs (Desveaux et al., 2005;Dubos et al., 2010;Seo et al., 2015)

The WRKY family is well known to mediate defense in plant in response to biotic and abiotic stresses but also are involved in other processes such as senescence, seed dormancy and development (Robatzek and Somssich, 2002;Rushton et al., 2010;Liu et al., 2016). The WRKY TF family has been well characterized in *A. thaliana*, comprising 74 members. Most of them are responsive to pathogen infection or signal molecules (Pandey and Somssich, 2009), modulating either positively or negatively the plant defense responses (Eulgem and Somssich, 2007). Many WRKYs have been described as positive regulators of SA-dependent responses in *Arabidopsis* (AtWRKY18, AtWRKY38, AtWRKY53, AtWRKY54, AtWRKY58, AtWRKY59, AtWRKY66, and AtWRKY70), being up-regulated during the NPR1-dependent SAR activation (Wang et al., 2006;Ishihama and Yoshioka, 2012). However, WRKY TFs often exhibit a dual activity in plant defense, depending on the type of pathogen. For instance, AtWRKY70 plays an important role as integrating signals from SA- and JA- dependent response, being responsible for inducing SA-responsive PR genes to enhance the resistance to biotrophic pathogens, at the time that repress the expression of JA-responsive genes, compromising resistance to necrotrophs in vegetative tissue (Li et al., 2004;Li et al., 2006). Similarly, AtWRKY50 and AtWRKY51 act as positive regulators of SA-mediated signaling, as well as negative regulators of JA-mediated signaling in Arabidopsis leaves (Gao et al., 2011). Also in *Arabidopsis* vegetative tissues, AtWRKY3 and AtWRKY4, two structurally similar WRKYs, have been described as positive regulators of plant resistance to necrotrophic pathogens such as Botrytis cinerea, but AtWRKY4 negatively affects the resistance to biotrophic pathogens such as *Pseudomonas syringae* (Lai et al., 2008). Also, constitutive expression of *AtWRKY33* conferred increased resistance to fungal necrotrophic pathogens, but enhanced susceptibility to the bacterial pathogen *P. syringae* (Zheng et al., 2006). On the contrary, the overexpression of *VvWRKY52* in *A. thaliana* green tissues enhanced resistance to biotrophic fungi *Erysiphe cichoracearum* and *P. syringae* pv. tomato DC3000 (Pst DC3000), but increased susceptibility to the necrotrophic pathogen *B. cinerea* (Wang et al., 2017).

Although many members of the WRKY gene family have been extensively studied in model plants using vegetative tissues, little is known about their defense-related function and regulation in strawberry, particularly in fruit. Strawberry fruit ripening changes the hormonal balance over time of auxins, ABA and gibberellins, among others (Symons et al., 2012), with potential crosstalk effects on the main SA- and JA/ET- defense pathways (Pieterse et al., 2012) and differential expression of genes

involved in both constitutive and induced defense mechanisms (Guidarelli et al., 2011). To date, 33 out of 62 FvWRKY transcription factors has been reported to be differentially regulated in the wild diploid strawberry species *Fragaria vesca* in response to powdery mildew infection (Wei et al., 2016).

FaWRKY1 was the first strawberry WRKY TF identified as mediator of defense response against to *C. acutatum* in cultivated strawberry (Encinas-Villarejo et al., 2009). *FaWRKY1* encodes an AtWRKY75-*like* transcription factor type IIc, which is up-regulated after *Colletotrichum* infection and responds to defense-related hormones such as SA, JA, ABA and wounding, being its expression dependent on strawberry cultivar and tissue (Encinas-Villarejo et al., 2009;Amil-Ruiz et al., 2016). In an attempt to characterize the function of this gene within the strawberry defense mechanism we previously undertook the heterologous overexpression of the FaWRKY1 in *A. thaliana Atwrky75* mutant and wild type (Encinas-Villarejo et al., 2009). The overexpression of *FaWRKY1* in the *Atwrky75* insertional mutant reverted the enhanced susceptible phenotype of the mutant, and even increased resistance over the wild type to avirulent strains of Pst DC3000. This resistant phenotype was associated with a strong oxidative burst and glutathione-S-transferase (GST) induction and was uncoupled to pathogenesis-related (PR) gene expression. These results proved for the first time a role of *FaWRKY1* gene in defense response and demonstrated that this strawberry gene could act as a positive regulator of resistance during compatible and incompatible interactions of a gram-negative phytopathogenic bacteria, in a heterologous plant system, pointing out a relevant role of this gene in the defense mechanism of strawberry. However, in strawberry, high level of *FaWRKY1* expression positively correlated with high degree of fruit infection by *C. acutatum* (Encinas-Villarejo et al., 2009).The observed differences between Arabidopsis and strawberry could reflect a distinctive modulation of FaWRKY1 biological function in different plant species and/or plant tissues or against different phytopathogens. Recently, the WRKY75 orthologs in apple and a rose were found upregulated in leaves in response to *Alternaria alternata* (Zhu et al., 2017), *Podosphaera pannosa* and *Diplocarpon rosae* (Neu et al., 2019), respectively, pointing out a role of the WRKY75-like TFs in defense responses on Rosaceae species, irrespectively of the pathogen's lifestyle. It is worthwhile to note that *AtWRKY75*-like genes also have been described to act differently according to the pathogen lifestyle. Thus, overexpression of VvWRKY1 increased the resistance of grapevine to the biotrophic pathogen Plasmopara viticola (Marchive et al., 2013). However, the GbWRKY1 acted as a negative regulator of the JA-mediated defense response in cotton and the silencing of this gene resulted in increased resistance to the necrotrophic *B. cinerea* and the hemibiotrophic *Verticilium dahliae* (Li et al., 2014).

To get further insight into the biological role that *FaWRKY1* gene plays within the mechanism of resistance to pathogens in strawberry, we have transiently silenced and overexpressed this gene in strawberry fruit using a modified *Agrobacterium*-mediated transient transformation methodology. Results demonstrate that susceptibility to *C. acutatum*, is significantly reduced in strawberry fruit where *FaWRKY1* was transiently

silenced whereas its ectopic overexpression does not substantially change susceptibility to this pathogen. Our study unravels a biologically relevant function of FaWRKY1 as negative regulator of resistance to *C. acutatum* infection in strawberry fruit and contrast with the previous role as a positive regulator of resistance found after its ectopic expression in *Arabidopsis*. Also, the expression pattern of some potential FaWRKY1 target genes was analyzed. Taken together, results shed light into the intricate FaWRKY1 regulatory network of strawberry fruit defense response against *C. acutatum*.

4.3. Materials and Methods

4.3.1. Fungal and Plant Material

Colletotrichum acutatum isolate CECT 20240 was maintained on potato dextrose agar (Duchefa) at 20 °C with 16/8 photoperiod or grown in strawberry agar (500 g/L grinded strawberry red fruits, 1.5% bacteriological agar) to increase the infectivity prior to pathogen inoculation. Stock conidia suspensions (106 conidia/mL) were prepared by scraping the surface of four-week old mycelia, in sterile distilled water containing 0.03% (v/v) Tween-80, then filtered in glass wool previously to the quantification the conidia concentration with a Neubauer Chamber Cell Counting. For pathogen inoculations, diluted 105 conidia/mL suspensions were prepared. Strawberry fruits (*Fragaria × ananassa* cv. Primoris) were grown under field conditions in Huelva (Finca Experimental “El Cebollar”, IFAPA), in southwestern Spain.

4.3.2. Plasmid construction for strawberry fruit transformation

All amplified sequences and specific primers used for plasmid constructs are described in Table S1. Binary plasmids (pK7WG2.0 and pKGWFS7.0) were obtained from VIB Plant Systems Biology (Belgium). pFRN binary vector was courtesy of Dr. Marten Denekamp, Department of Molecular Cell Biology, University of Utrecht (The Netherlands). For all the cloning steps using gateway technology, standard Invitrogen protocols were used.

For the transient overexpression of the *FaWRKY1* gene in strawberry fruits, the plasmid pK7WG2::FaWRKY1 (35S::FaWRKY1) previously described in (Encinas-Villarejo et al., 2009) was used. For the spatial localization and time course visualization of the transgene after agroinfiltration in strawberry fruits following this innovative procedure, a 1035 bp DNA fragment carrying the complete CaMV35s promoter was specifically amplified from pK7WG2.0 vector with primers p35S-attBfw and p35S-attBrv and cloned into the pDNOR221 entry vector. This 1035 bp DNA fragment was later transferred to pKGWFS7.0 destination vector to obtain the pKGWFS7::pCaMV35s::GUS (p35S::GUS) plasmid derivative where the *β-glucuronidase* gene is driven under control of the *CaMV35* promoter.

Table S1. Primers and sequences used in this study.

Name	Description/ortolog	RT-qPCR primer sequence (5´ - 3´)		Observation
FaWRKY1	*WRKY DNA-binding protein 75* gene07210	sense chain anti-sense chain	ACAGCAGTAAGATTAGGGATGAAGAAGGGAG GCTTCTTCACATTGCAACCCTGATGCGTG	Amil-Ruiz et al.,2016 RT-qPCR
FaCHI4-2	Chitinase class IV gene02717	sense chain anti-sense chain	TGCCGGCAAGAGCTTCTACACTAGACAG TGTGAAGGATGTGTCGCAGTAGGTGG	
FaACTIN	*Actin* gene26612	sense chain anti-sense chain	GTATACATCCTGAAGTGGTAGACGGAGG GGGCCAGAAAGATGCTTATGTCGG	Amil-Ruiz et al.,2013 RT-qPCR
FaEF1a	*elongation factor 1-alpha* gene28639	sense chain anti-sense chain	TGGATTTGAGGGTGACAACATGA GTATACATCCTGAAGTGGTAGACGGAGG	
FaICS1	*Isochorismate syntase1* gene25950	sense chain anti-sense chain	TTGGTCAGTGTTGCTGGTGT GTAGCCAAAGGCCTCCCATT	This work RT-qPCR
FaCAT	*Catalase* gene10917	sense chain anti-sense chain	CCTGCCCTTATTGTCCCTGG GTGAGCAGACTTGGGAGCAT	
FaWHY1	*Whirly1* gene04012	sense chain anti-sense chain	TTTCGGAAGAGACGACTCGC ACATAAAACCTAGCCGGCCC	
FaWHY2	*Whirly2* gene31174	sense chain anti-sense chain	ACACAAAGCACCGGCTTTTC AGAGAGTGCAGCTTTGCCTT	
FaJAZ1	*FvTIFY10A* gene12541	sense chain anti-sense chain	TGGGAGATCTGAACCTCGTC TTCCTCGGTTTCTCCATCAC	Garrido-Bigotes et al.,2018 RT-qPCR
FaJAZ4	*FvTIFY6B* gene06180	sense chain anti-sense chain	AGAAGTGCTGGTGCACATTG TGGGCATAAATCTGGAGGAC	
FaJAZ5	*FvTIFY11-Like* gene05383	sense chain anti-sense chain	CACCATGAACTTGCTCAACG GAAAGGTCGCTGAAGACGAG	
FaJAZ7	*FvTIFY5B* gene24321	sense chain anti-sense chain	GGATGAGCAGACCAGACAGG AAACATAAACCCGGCCATCG	
FaJAZ8.1	*FvTIFY5A* gene30624	sense chain anti-sense chain	GAGGAGGAACTGCAATTTGG AAGAGGGAAGCCGGAATTAG	
FaJAZ9	*FvTIFY6B* gene09356	sense chain anti-sense chain	GGATGAGCAGACCAGACAGG AAACATAAACCCGGCCATCG	
FaJAZ10	*FvTIFY9* gene07265	sense chain anti-sense chain	TTCCAGAAGTTCCTCGAACG GATTTCCTGGCTGCAATCAC	
FaJAZ12	*FvTIFY3B* gene12975	sense chain anti-sense chain	GAAGCGTAGGGACAGATTGG AACCGGAAGAAGCATCATTG	
pFRN::FaWRKY1 (RNAi)	FaWRKY1 RNAi fragment amplification	sense chain anti-sense chain	ATGGATACCTACCCAGCATTCTA TCCCTTCTTCATCCCTAATC	Cloning of a 272 bp fragment from FaWRKY1 in pFRN
35S::FaWRKY1 (Overexpression)	FaWRKY1 full cDNA amplification	sense chain anti-sense chain	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAACAATG GATACCTACCCAGCATTCTA GGGGACCACTTTGTACAAGAAAGCTGGGTCTTACAAAG CACGCACAGCAGGCA	Encinas-Villarejo et al.,2009
pCaMV35s::GUS	Cauliflower mosaic virus (CaMV) promoter	sense chain anti-sense chain	GGGGACAAGTTTGTACAAAAAAGCAGGCTACTAGAGC CAAGCTGATCTCCTT GGGGACCACTTTGTACAAGAAAGCTGGGTTCGACTAGA ATAGTAAATTGTAATG	Cloning of CaMV promoter (1035 bp) in pKGWFS7.0
FaWRKY1-RNAi	*WRKY DNA-binding protein 75*	ATGGATACCTACCCAGCATTCTATTCTTCTTCATCAACACCACCTTCTGCTGCTG CTTCTTCGCTGTCATTGAACATGGTGAACACTCATCCTCATCATGCTTACGGTAA CGATCATCAGTACCAAGCTAGCAATAACAAGGAAAATGGGTTCTTGGGGCTGAT GTCAGAGATGGAGGTTTCGAACAGCATGAGTAGTATTACCCAGCAGAGTATGAA AAGCTTTGGGGAGGGTGAAAGTAATACAGCAGTAAGATTAGGGATGAAGAAGG		FaWRKY1 fragment (272) bp for RNAi gene silencing

The silencing FaWRKY1-RNAi cassette was constructed as follows: a 272 bp non-conserved region of this gene corresponding to the 91 first amino-acids was PCR amplified from pK7WG2::FaWRKY1 with specific WRKY1-RNAi forward and reverse primers and cloned into pCR8/GW/TOPO (Invitrogen) as an entry vector. The cloned fragment was subsequently transferred to the destination pFRN vector to obtain the RNAi silencing construct, pFRN::FaWRKY1-RNAi. The correct sense and antisense orientation of the 272 bp WRKY1 DNA fragment spaced by the CHS intron was confirmed by sequencing prior further manipulations. All these constructs, including their corresponding empty vectors, were introduced into *Agrobacterium tumefaciens* strain AGL0 (Lazo et al., 1991) using the freeze–thaw shock method (Holsters et al., 1978). *A. tumefaciens* strains were grown at 28 °C in Luria–Bertani (LB) medium with appropriate antibiotics. When the culture reached an optical density of about 0.8 at 600 nm (OD600), cells were harvested and resuspended in a modified MacConkey agar (MMA) medium (Spolaore et al., 2001). After 1 h of incubation at 22°C in dark, *Agrobacterium* suspensions were injected into fruits using one-milliliter syringes.

4.3.3. Fruit Agroinfiltration and Experimental design

Strawberry fruits at pink/turning stage were collected with a pigmentation degree of approximately 25% (Aharoni et al., 2002). All fruits were excised along with their pedicels, of ten centimeters long, then sterilized with commercial bleach (1:60 v/v) and cultured by pedicel immersion in sterile rich medium MS (0.25x Murashige Skoog and 4 g sucrose per liter). Every strawberry fruit was maintained in this medium for all assay period (6 days), with new fresh medium changes, every two days. A modified protocol of agroinfiltration previously described (Spolaore et al., 2001;Hoffmann et al., 2006) was performed to reduce variability among fruits and be able to compare the defense response to pathogen inoculations between halves of the same fruit (Figure S1). Thus, a half of the fruit was infiltrated with *Agrobacterium* bearing the query transgene construct and the opposite half with *Agrobacterium* bearing the corresponding empty vector, as a control. In order to clearly distinguish both fruit halves for later manipulations, two sepals were removed from the half corresponding to the query constructs (either silencing or overexpression) before agroinfiltration. Short needles (23Gx25 mm) were employed for infiltration to ensure that each *Agrobacterium* suspension remains in the corresponding fruit half and did not spread within the opposite half of fruit. The agroinfiltration was carried out in the center of every strawberry fruit half, approximately. Although 1 ml of *Agrobacterium* suspension was infiltrated in most of strawberry fruit halves, this volume was slightly adjusted according to the size of the fruit until complete run off. Two days after the agroinfiltration, both halves of the fruit were inoculated with *C. acutatum* using 5 mm paper discs embedded in a 10^5 conidia/mL suspension. The embedded discs were placed on the strawberry surface mainly located on the agroinfiltration point. A subset of the agroinfiltrated fruits was reserved as “non-infected fruits” and was not inoculated with the pathogen. All fruits were stored in closed chamber with 75-80%

humidity at 25ºC for five days. Three fruits were collected every day and samples from each of the two halves of the collected fruits were immediately frozen in liquid nitrogen and transferred to -80 °C until use.

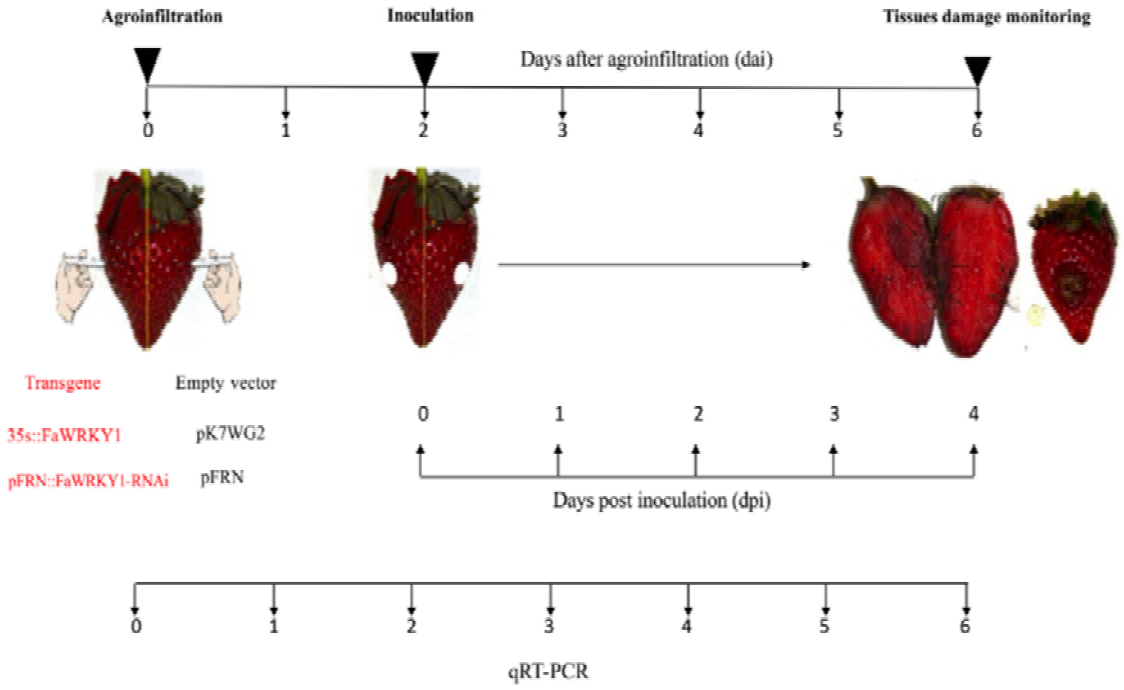


Figure S1. Experimental procedure for tissue damage assessment produced by *C. acutatum* in strawberry fruits, agroinfiltrated with AGL0 bearing overexpression or silencing *FaWRKY1* gene constructs.

Healthy fruits grown under field conditions were selected and used following our experimental design, which was two times repeated during strawberry fruiting season for two years (288 fruits for the silencing or overexpression, including its corresponding control). The total number (576) of strawberry fruits used in this work is summarized in Table S2. For silencing of FaWRKY1 gene, 144 out of 288 fruits were agroinfiltrated with the pFRN::FaWRKY1-RNAi construct in one fruit half and with the corresponding empty vector in the opposite fruit half. To analyze the biological effect of the agroinfiltration process in the expression of the endogenous *FaWRKY1* gene, 144 fruits were agroinfiltrated only in one fruit half with the empty vector but no agroinfiltration was accomplished in the corresponding opposite fruit half. In every case, 120 out of the 144 fruits were inoculated with the pathogen in both fruit halves (24 out of 120 fruits were used for expression studies and 96 remaining fruits for statistical purposes), leaving 24 fruits without inoculation as control, in order to analyze the silencing of *FaWRKY1* upon no infection condition. From each experimental condition, three fruits were collected every 24 h for 6 days, which were

used for the *FaWRKY1* gene expression analysis by RT-qPCR. A similar number of fruits and identical protocol was followed for the *FaWRKY1* overexpression experiments but plasmid 35S::FaWRKY1 was used instead.

Table S2. Distribution of fruit number according to season and constructs used in the present study

		1° year		2° year		Total	
Construct	**Purposes**	**+I**	**-I**	**+I**	**-I**	**+I**	**-I**
pFRN::FaWRKY1-RNAi	Expression	24	24	-	-	120	24
	Statistic	48	-	48	-		
pFRN (RNAi control)	Expression	24	24	-	-	120	24
	Statistic	48	-	48	-		
	Total per year	192		96		288	
35S::FaWRKY1	Expression	24	24	-	-	120	24
	Statistic	48	-	48	-		
pK7WG2 (OE control)	Expression	24	24	-	-	120	24
	Statistic	48	-	48	-		
	Total per year	192		96		288	

4.3.4. Histochemical GUS Assay

A total of 30 strawberry transient transform fruit halves (three fruits were collected every 24 h for 7 days) were used for the histochemical assay. GUS activities in strawberry was performed as described by (Jefferson et al., 1987) using a modified staining solution following the manufacturer (Gold Biotechnology) instructions containing: 2 mM X-gluc in 100 mM sodium phosphate buffer (pH 7.5), 10 mM EDTA, 0.1% (v/v) Triton X-100, 1.0 mM potassium ferricyanide.

4.3.5. Total RNA Extraction and Real-Time qPCR

Total RNA from strawberry tissues was isolated as described previously (Casado-Díaz et al., 2006), treated with DnaseI (Invitrogen) for residual DNA removal, and further purified with the RNeasy MinElute Cleanup Kit (QIAGEN). Purified RNA was quantified by NanoDrop 1000 Spectrophotometer (Thermo scientific). RNA integrity was checked using the Agilent 2100 Bioanalyzer (Agilent Technologies, Deutschland). First-strand cDNA synthesis was carried out using 1 µg of purified total RNA as template for a 20 µL reaction [iScript cDNA Synthesis kit (Bio-Rad)]. RT reactions were diluted 5-fold with nuclease-free water prior to RT-qPCR.

Specific primer pairs set were designed using Oligo Primer Analysis software version 6.65, tested by dissociation curve analysis, and verified for absence of non-specific amplification. The expression levels were calculated according to the 2-ΔΔCT method (Livak and Schmittgen, 2001) and, normalized according to two housekeeping gene *actine 1* (*FaACT1*) and *elongation factor 1α* (*FaEF1a*) (Amil-Ruiz et al., 2013). RT-qPCR runs were performed using specific primers (Table S1) in two technical replicates in the same run and three biological replicates in different runs, as described previously (Encinas-Villarejo et al., 2009), using SsoAdvanced™ SYBR® Green Supermix, and MyIQ v1.004 and iCycler v3.1 real-time PCR systems (Bio-Rad). Mean PCR efficiencies were calculated by LinRegPCR software (Ruijter et al., 2009). All RT-qPCR primers used in this study have similar PCR efficiencies.

The level of silencing and overexpression of *FaWRKY1* was calculated for each time and normalized as the relative expression value of this gene between the agroinfiltrated fruit half with the query cassette construct and the corresponding agroinfiltrated opposite fruit half with the control vector.

4.3.6. Tissue Fruit damage evaluation and Statistical Analysis

Seventy fruits (n=70) for the silencing and sixty fruit (n=60) for the overexpression experiment were phenotypically observed and evaluated for tissue damage at 4 days post inoculation with *C. acutatum*. The severity of tissue damage was carried out on 5-scale according to (Jin Choi et al., 2016). Essentially, 1, symptomless tissues (0% fruit halve damaged); 2, weakly visible lesion (up to 10% fruit halve damaged); 3, moderate

lesion (10-25% fruit halve damaged); 4, enlarged lesion (25-50% fruit halve damaged); 5, very affected fruit (> 50% fruit halve damaged). Two different ratios were calculated: internal damage ratio and external damage ratio, both resulting from dividing the internal or external tissue damage value of the fruit half where the transgene was overexpressed or silenced by the tissue damage value corresponding to the opposite half of the same fruit infiltrated with *Agrobacterium* bearing the empty vector. Fruits where both halves were infiltrated with *Agrobacterium* bearing the empty vector were used as control for statistical purposes. Means and SE were obtained by Fisher's LSD test (α=0.05) by Statistix software (v9.0). A ratio of 1, clearly indicate no differences between both halves of the same fruit.

Real Time-qPCR data were statistically analyzed in Microsoft Excel, using the Real Statistics Resource Pack software, release 5.4 (http://www.real-statistics.com/). All data were tested for normality using Shapiro-Wilk test (α=0.05). One-way ANOVA, followed by Dunnett's test or Tukey´s test post-hoc were performed at α=0.05 and 0.01. Three biological replicates were used (n=3).

4.4. Results

4.4.1. Spatial-temporal expression analyses of the transgene after *Agrobacterium* infiltration in strawberry fruits

To validate the methodology and to identify the spatio-temporal gene expression of the transgene after strawberry fruit *Agrobacterium* infiltration (agroinfiltration) in our experimental conditions, β-glucuronidase (GUS) activity was determined in fruits in which only one half was agroinfiltrated with the construct p35S::GUS (Figure 1). GUS activity was monitored in longitudinal sections of these strawberry fruits every 24 hours, up to seven days after agroinfiltration (7 dai). As shown in Figure 1, GUS activity became clearly visible at the second day after agroinfiltration (2 dai), and this expression was detected only within the agroinfiltrated fruit half. Interestingly, GUS activity slightly increased at 3 dai and it was maintained up to 7 dai and was confined only within the agroinfiltrated fruit half being limited by the pith. Therefore, no activity was clearly visible in the opposite half of fruit and aloof regions of injection point after seven days (Figure 1). Considering these data, two days after agroinfiltration (2 dai) was chosen as the appropriate time to make the C. acutatum inoculations in those experiments designed to test loss and gain of FaWRKY1 function by Agrobacterium transient expression in fruit.

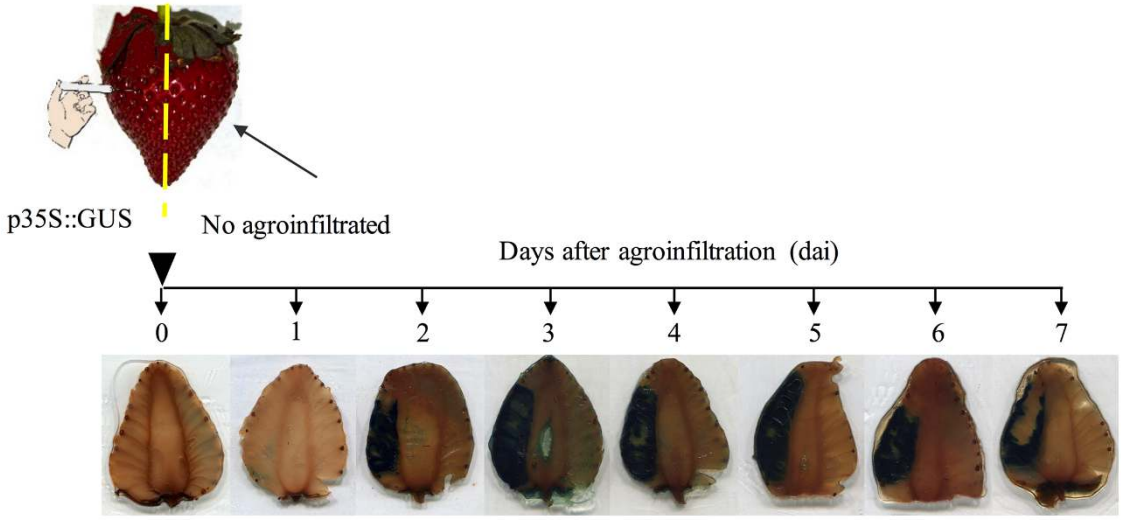


Figure 1. Spatial and time course expression analysis of GUS reporter gene in strawberry fruit (Fragaria x ananassa) agroinfiltrated with pCaMV35s::GUS-GFP (p35S::GUS). Histochemical GUS staining was performed in longitudinal sections of fruits which only one half of the fruit was agroinfiltrated with *A. tumefaciens* carrying the plasmid pCaMV35s::GUS-GFP. The GUS activity was determined every 24h, up to seven days after agroinfiltration, as described in materials and methods. The agroinfiltrated fruit half was marked and easily distinguished by removing two sepals before infiltration and it is shown as the left side of every fruit slice in this picture.

4.4.2. Changes in the expression pattern of *FaWRKY1* after *Agrobacterium* infiltration and *C. acutatum* inoculation are independently distinguished in strawberry fruit.

A time course analysis by real-time PCR was performed on strawberry fruits after agroinfiltration and *C. acutatum* inoculation. Thus, fruits were agroinfiltrated in one half with the empty pFRN vector. Two days after agroinfitration, 24 out of the 48 fruits were also inoculated in both halves with *C. acutatum*, while the other 24 fruits remained uninfected (treated with mock-soaked discs), as control reference. In order to distinguish changes in the *FaWRKY1* expression pattern only due to a response of the fruit to A. tumefaciens infiltration and/or *C. acutatum* infection, the expression pattern of *FaWRKY1* was determined over time by comparing both, fruit halves agroinfiltrated and inoculated versus fruit halves agroinfiltrated and non-inoculated, and fruit halves only inoculated versus non-inoculated ones (Figure 2).

Results in Figure 2A show the complexity of the *FaWRKY1* expression pattern under all conditions tested. When data from fruits inoculated with *C. acutatum* was normalized to non-inoculated ones, a significant accumulation of *FaWRKY1* transcripts were detected in response to this pathogen inoculation, which reached a peak level at 5 dai, corresponding to 3 days post inoculation (3 dpi) (Figure 2B). This expression decreased to lower level at 6 dai (4 dpi). Moreover, the expression pattern of *FaWRKY1* was more complex when data from agroinfiltrated and inoculated fruits was normalized to the corresponding agroinfiltrated and non-inoculated ones (Figure 2C). Thus, *FaWRKY1* was significantly induced in response to *Agrobacterium* infiltration (Figure 2C) and this expression reached its peak at 2 days after infiltration (2 dai), but decreased to lower levels at 3 dai. In addition, a fast increase in *FaWRKY1* transcript accumulation was later detected, reaching peak levels at 5 dai and quickly decreasing at 6 dai. Interestingly, this second peak of expression correlated to that showed in Figure 2B, thus indicating that *C. acutatum* was able to induce *FaWRKY1* expression in the agroinfiltrated fruits in a similar way to that observed in the non-agroinfiltrated fruits. These results state the expression pattern of FaWRKY1 gene in strawberry fruit in our experimental system after *Agrobacterium* infiltration and *C. acutatum* inoculation and indicate that changes in the expression pattern of this gene due to each event are independently distributed and can be distinguished over time (Figure 2C).

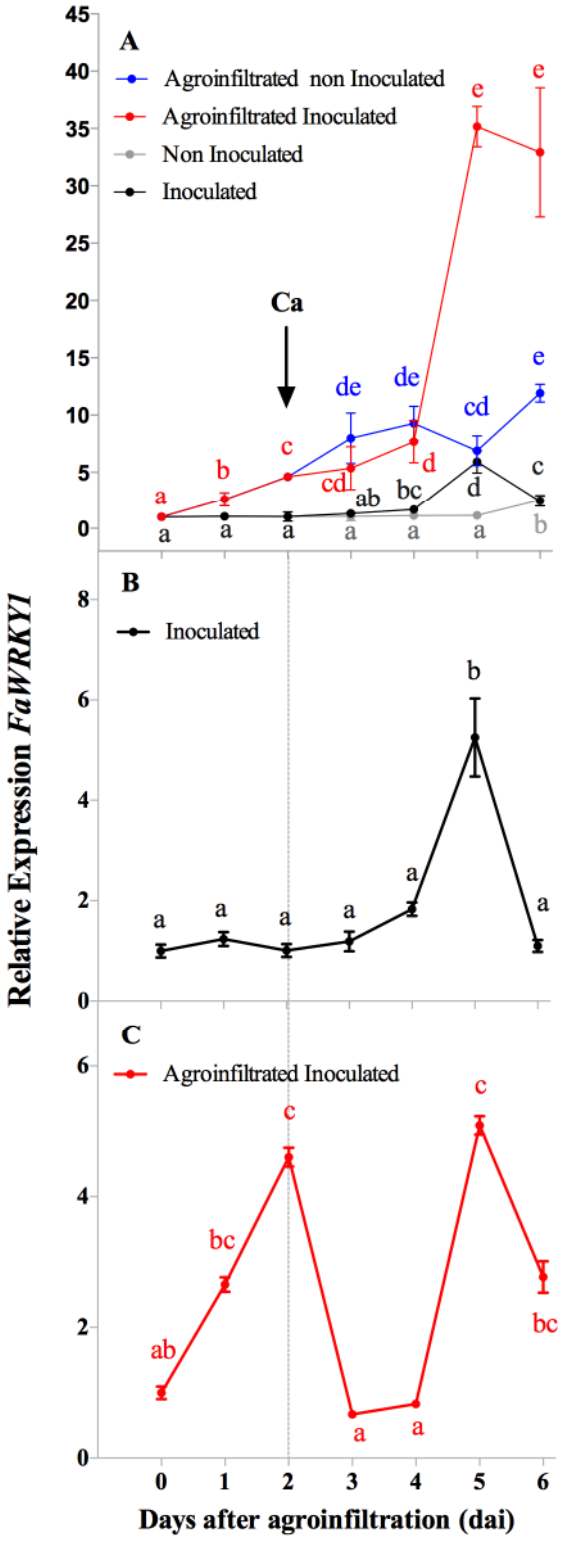


Figure 2. Gene expression pattern of *FaWRKY1* in strawberry fruit after *A. tumefaciens* infiltration and *C. acutatum* inoculation. To easy follow, time scale is represented here along 6 days after fruit infiltration with *A. tumefaciens* bearing the pFRN vector even though in A and B, non agroinfiltrated fruit are represented. The arrows indicate the time of inoculation with *C. acutatum* (Ca). In A, B, and C, the relative expression of *FaWRKY1* is represented with respect to time zero; (A) relative expression of *FaWRKY1* in agroinfiltrated or non-agroinfiltrated strawberry fruit upon *C. acutatum* infection or mock; (B) the relative expression of *FaWRKY1* in non-agroinfiltrated and inoculated fruit was normalized with respect to the non-agroinfiltrated non-inoculated ones; (C), the *FaWRKY1* expression in pFRN agroinfiltrated and inoculated fruit was normalized with respect to the agroinfiltrated and non-inoculated ones. Values are the means of three biological replicates. Means followed by the same letters in each trait are not significantly different (α=0.05), according to Tukey's test.

4.4.3. The transient silencing of *FaWRKY1* gene in strawberry fruit reduced fruit tissue damage after *C. acutatum* inoculation

Strawberry fruits were independently agroinfiltrated on opposite halves of the same fruit with silencing (pFRN::FaWRKY1-RNAi) and empty vector (pFRN) constructs. Two days after agroinfitration, 24 out of the 48 fruits were also inoculated in both halves with *C. acutatum*, while the other 24 fruits remained uninfected (treated with mock-soacked discs) (see experimental design details in Figure S1).

Changes in the expression pattern of *FaWRKY1* were analyzed in agroinfiltrated fruit halves, which were either not exposed (Figure S2A) or exposed (Figure S2B) to *C. acutatum* infection. Results in Figure S2A show the induction pattern of the *FaWRKY1* gene over time, due to the agroinfiltration event in both pFRN and pFRN::FaWRKY1-RNAi agroinfiltrated fruit halves. This induction was always much lower in pFRN::FaWRKY1-RNAi silenced fruit halves compared to that of pFRN ones. Increases in *FaWRKY1* transcript accumulation after *C. acutatum* inoculation were also detected in fruit halves which were previously and independently agroinfiltrated with empty vector (pFRN) and pFRN::FaWRKY1-RNAi silencing construct, respectively (Figure S2B). Also, a much lower level of *FaWRKY1* expression was detected along all time points for pFRN::FaWRKY1-RNAi fruit halves compared to pFRN ones. This reduction in the *FaWRKY1* expression was markedly relevant at 5 dai (corresponding to 3 dpi), where an increasing value of 35-fold was detected for the expression of *FaWRKY1* gene in pFRN fruit halves whereas less than 9-fold increase was only detected in fruit halves agroinfiltrated with pFRN::FaWRKY1-RNAi.

The level of *FaWRKY1* gene silencing was calculated for every time and condition by normalizing the silenced fruit halves values to the control ones of the corresponding opposite fruit halves, and it is also represented in Figure S2A and S2B. A reduction of *FaWRKY1* transcript accumulation was found as early as two days after agroinfiltration and remarkable silencing values were observed at 4 dai (2 dpi), 5 dai (3dpi), and 6 dai (4 dpi) in both inoculated and non-inoculated fruit. These results clearly indicated that in our experimental system the *FaWRKY1* gene is successfully silenced in strawberry fruit after agroinfiltration with the pFRN::FaWRKY1-RNAi silencing construct.

Taking into account all these results, the evaluation of fruit tissue damage and the comparative analysis of susceptibility to *C. acutatum* between the two halves of the same fruit (one agroinfiltrated with the silencing construct and the other with the empty vector as control) were accomplished after 6 dai (4 dpi), in a total of 70 fruits (Figure 3). In general, no relevant differences in the external tissue damage were visually observed in both opposite halves of the same fruit. Thus, mycelial growth, accompanied by tissue browning and depressed necrosis, was clearly visible surrounding the inoculation area after 4 dpi (Figure 3A). However, when the internal tissue damage was evaluated, a relevant reduction was clearly detected within fruit halves agroinfiltrated with the silencing construct compared to the corresponding opposite fruit halves agroinfiltrated with the empty pFRN vector (Figure 3B).

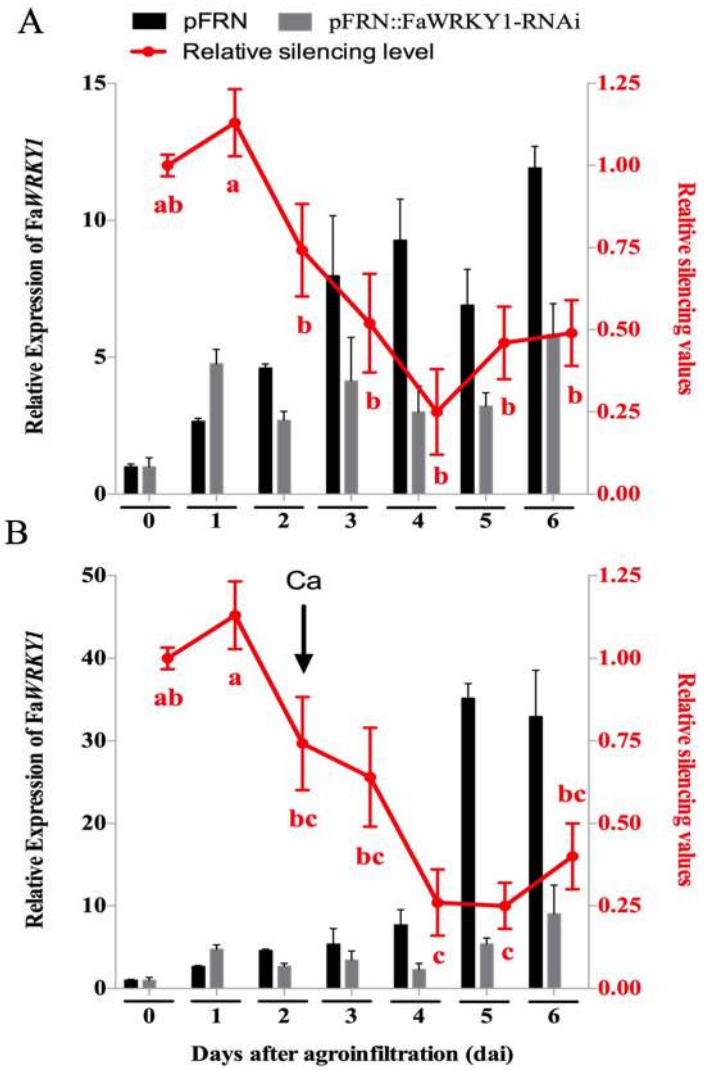


Figure S2. Effect of transient silencing of *FaWRKY1* gene in strawberry fruit by agroinfiltration and *C. acutatum* infection. Gene expression kinetic analysis of *FaWRKY1* by RT-qPCR in each of the two strawberry fruit halves agroinfiltrated with pFRN (black line) and pFRN::FaWRKY1-RNAi (grey line), respectively, and non-inoculated (A) or inoculated with C. acutatum (B). Data were normalized with respect to the transcript level of the housekeeping elongation factor 1α and actine 1 genes, and scored during six days. Values are the means of three biological replicates. In red, the level of *FaWRKY1* silencing was calculated as a ratio value between the gene expression values found within the pFRN::FaWRKY1 fruit half with respect to the corresponding pFRN half. Means followed by the same letters in each trait are not significantly different at 5% level, according to Tukey's test.

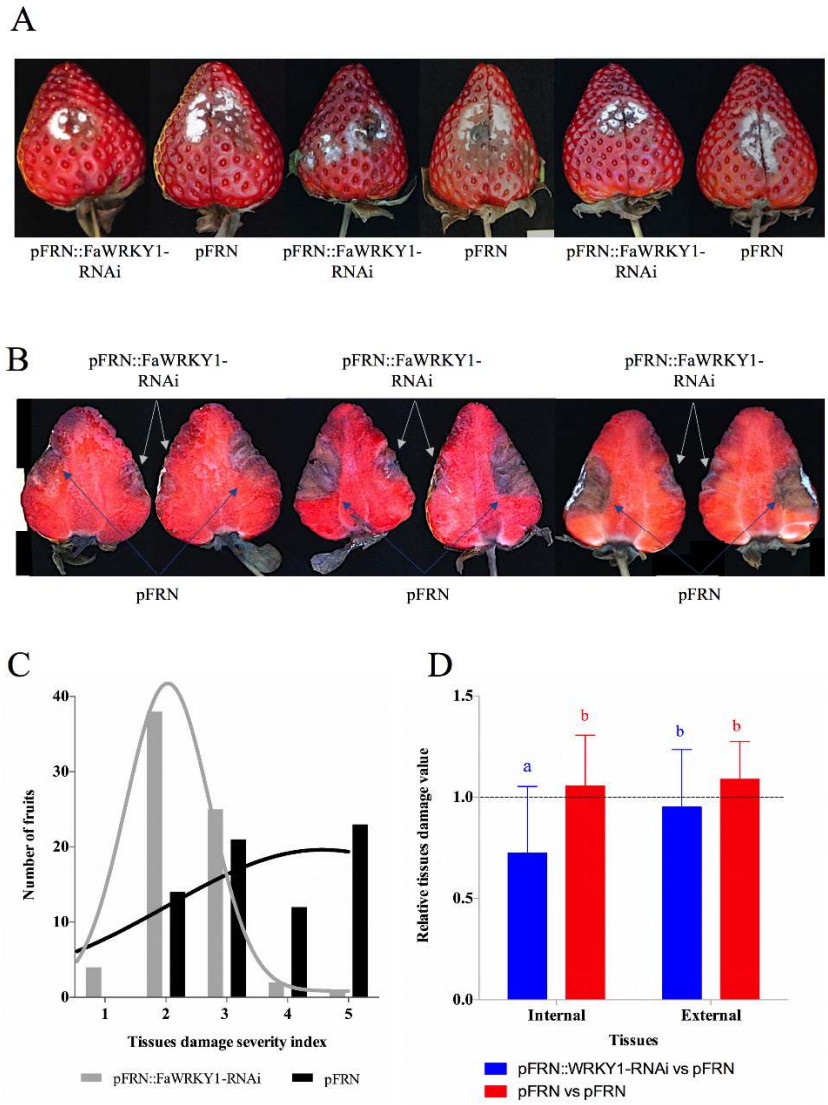


Figure 3. The silencing of *FaWRKY1* in strawberry fruit enhances resistance to *C. acutatum* infection. (A) External surface disease symptoms on the two agroinfiltrated opposite halves (pFRN::FaWRKY1-RNAi and pFRN) of the same fruit. Three different fruits are shown as an example. (B) Internal tissue damage, of the same three fruits shown in A; white and blue arrows indicated the tissue area affected in the pFRN::FaWRKY1-RNAi or pFRN empty vector agroinfiltrated fruit half, respectively, at 4 dpi. (C) Distribution of a total of 70 strawberry fruits based on a 1 to 5 scale used to asses tissue damage in each of the two opposite fruit halves of the same fruit (1, no symptoms; 2, weakly visible lesion; 3, moderate lesion; 4, enlarged lesions; 5, very affected); black and grey bars, number of pFRN and pFRN::FaWRKY1-RNAi agroinfiltrated fruit halves, respectively, exhibiting the indicated grade of tissue damage at 4 dpi with *C. acutatum*; black and grey lines, indicate the distribution of each fruit half according to the severity scale. (D) Statistically analysis of internal and external tissue damage ratio of the two opposite halves of the same fruit, according to the 1 to 5 severity scale; blue and red bars, pFRN::WRKY1-RNAi/pFRN and pFRN/pFRN agroinfiltrated values, respectively. Data correspond to mean ± SD. Within each bars, means with different letters are significantly different by LSD test at $p < 0.05$. A ratio value of 1 indicate no differences between opposite halves of the same fruit.

Also, the distribution of damaged fruit, based on a 1 to 5 scale, was evaluated (Figure 3C). Hence, approximately 60% of the pFRN::FaWRKY1-RNAi agroinfiltrated fruit halves presented none or very small tissue damage (scores 1, and 2) , while 36% showed moderate damage (score 3), and only 4% presented a very large affected tissue area (scores 4, and 5). On the other hand, only 20% and 30% of the pFRN agroinfiltrated fruit halves showed small (score 2) or moderate (score 3) tissue damage, respectively, while a higher percentage of up to 50% were strongly affected by *C. acutatum* infection (scores 4, and 5) (Figure 3C).

A statistical analysis of the internal and external fruit tissue damage was conducted and is shown in Figure 3D. For internal fruit tissue damage, the values obtained by normalizing fruit halves transformed with the silencing construct with respect to the corresponding opposite fruit halves transformed with the empty pFRN vector were significantly reduced (mean value of 0,7257) compared to those obtained when both fruit halves were transformed with pFRN control constructs (mean value 1,0556). On the contrary, for external fruit tissue damage, all the ratio values show no significant differences. These results establish a positive correlation between the silencing of the FaWRKY1 gene and an increase of fruit resistance to *C. acutatum* (Figure 3D).

4.4.4. The transient overexpression of *FaWRKY1* in strawberry fruit did not alter susceptibility to *C. acutatum*

The expression pattern of *FaWRKY1* gene was analyzed in inoculated fruits in which one half was agroinfiltrated with the 35S::FaWRKY1 overexpression construct, and the opposite half with the empty pK7WG2 vector as control (Figure S3). Based on the results observed in Figure S2 for the silencing experiment, the expression of this transgene was expected to be highly induced at 4 and 5 dai. Accordingly, the *FaWRKY1* gene was highly induced, at both the fourth and fifth day after agroinfiltration in fruit halves transformed with the 35S::FaWRKY1 construct, in comparison to that of the corresponding opposites fruit halves both transformed with the pK7WG2 control vector.

The level of *FaWRKY1* overexpression was also calculated for every time by normalizing the overexpressed fruit halves values to that of the control ones of the corresponding opposite fruit halves, and it is also represented in Figure S3 (red line). Thus, increases of *FaWRKY1* transcript level of 20-fold and 10-fold that of control were found at 4 dai and 5 dai, respectively, in fruit halves transformed with the overexpression vector. These results clearly indicated that in our experimental system the *FaWRKY1* gene is successfully overexpressed in strawberry fruit after agroinfiltration with the 35S::FaWRKY1 construct. The evaluation of fruit tissue damage and the comparative analysis of susceptibility to *C. acutatum* between the two opposite halves of the same fruit (one agroinfiltrated with the 35s::FaWRKY1 construct and the other with the pK7GW2.0 empty vector as control) were accomplished after 6 dai (4 dpi), in a total of 60 fruits (Figure 4), as previously described for the silencing

experiments. Again, no relevant differences in the external tissue damage were visually observed between opposite halves (Figure 4A). When the internal tissue damage was evaluated, no relevant differences were also visually detected between fruit halves transformed with the 35S::FaWRKY1 construct compared to the corresponding opposite fruit halves transformed with the empty pK7GW2.0 vector (Figure 4B). In addition, the distribution pattern of damaged fruit was similar for the two fruit halves, irrespectively of the agroinfiltrated construct (Figure 4C). Interestingly, damage-free fruit halves were not observed and only few fruits (10% and 12% of the overexpressed and control fruit halves, respectively) showed small damaged region (score 2). Instead, most of fruit halves showed moderated damage (score 3), and a similar high percentage of fruit showed large or very large tissue damage (scores 4 and 5) on both fruit halves. In fact, the statistical analysis of internal and external fruit tissue damage did not show any significant difference, irrespectively of the agroinfiltrated construct (Figure 4D). These results indicate that the transient overexpression of *FaWRKY1* in strawberry fruit does not seem to substantially affect susceptibility to *C. acutatum.*

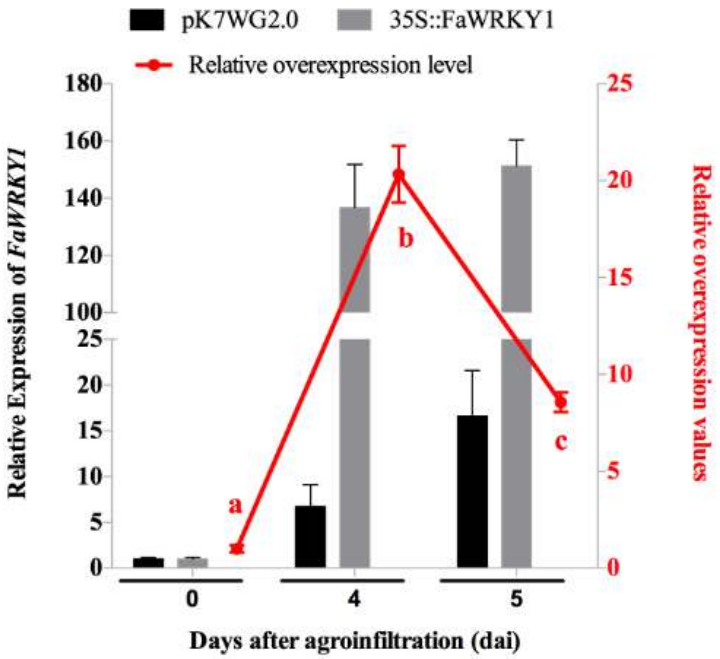


Figure S3. Effect of transient ectopic expression of *FaWRKY1* gene in strawberry fruit by agroinfiltration and *C. acutatum* infection. Gene expression kinetic analysis of *FaWRKY1* by RT-qPCR in each of the two strawberry fruit halves agroinfiltrated with pK7::WRKY1 (35s::FaWRKY1) (grey line) or pK7GW2.0 (black line), respectively, and inoculated with *C. acutatum*. In the graphics, standard value 1 at T0 was added to better illustrate changes. Data were normalized with respect to the transcript level of the housekeeping *elongation factor 1α* and *actine 1* genes. Values are the means of three biological replicates. In red, the level of overexpression of *FaWRKY1* was calculated over time as a ratio value between the gene expression values found within the pK7::WRKY1 fruit half with respect to the corresponding pK7WG2.0 half. Means followed by the same letters in each trait are not significantly different at 5% level, according to Tukey's test.

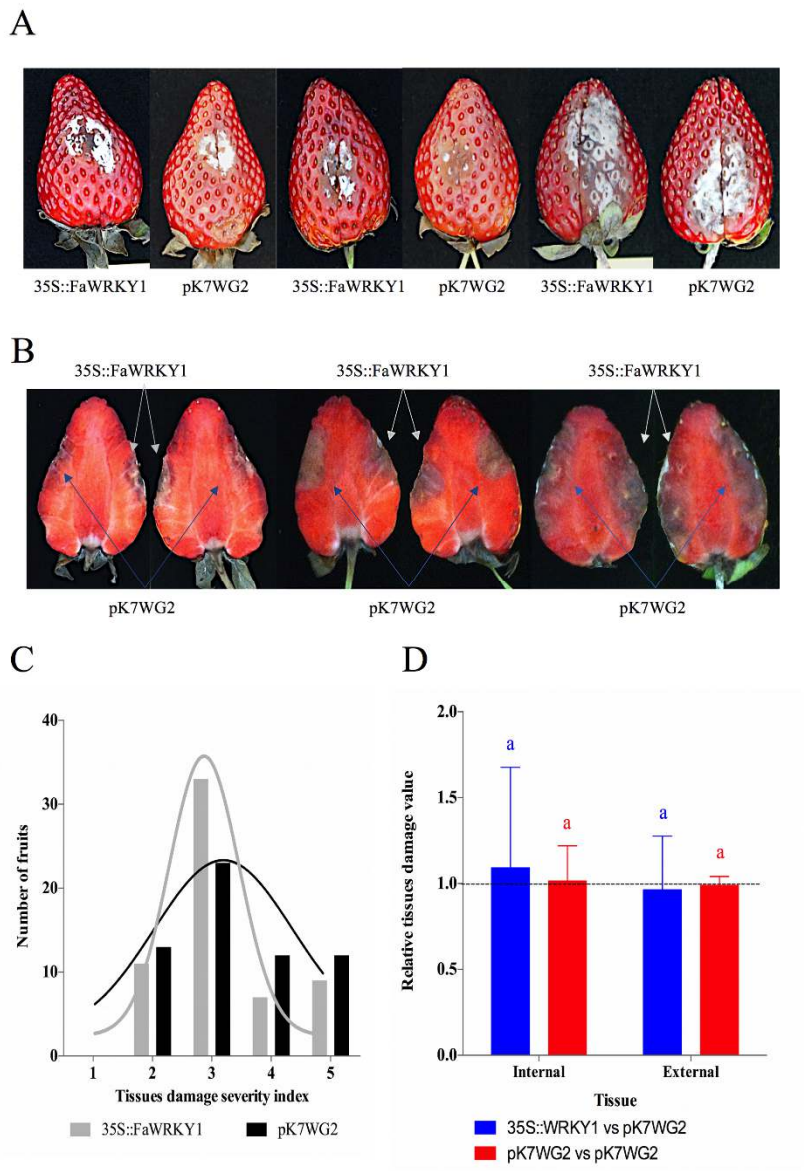


Figure 4. Transient overexpression of *FaWRKY1* in strawberry fruit does not substantially alter susceptibility to *C. acutatum* infection. (A) External surface disease symptoms on the two agroinfiltrated opposite halves (pK7WG2::FaWRKY1 and pK7WG2.0) of the same fruit. Three different fruits are shown as an example. (B) Internal tissue damage of the same three fruits shown in A; white and blue arrows indicates the tissue area affected in the pK7WG2::FaWRKY1 (35S::FaWRKY1) or pK7WG2.0 empty vector agroinfiltrated fruit half, respectively, at 4 dpi. (C) Distribution of a total of 60 strawberry fruits based on a 1 to 5 used to asses tissue damage in each of the two opposite fruit halves of the same fruit (1, no symptoms; 2, weakly visible lesion; 3, moderate lesion; 4, enlarged lesions; 5, very affected); black and grey bars, number of pK7WG2.0 and pK7WG2::FaWRKY1 agroinfiltrated fruit halves, respectively, exhibiting the indicated grade of tissue damage at 4 dpi with *C. acutatum*; black and grey lines, indicate the distribution of each fruit halves according to the severity scale. (D) Statistical analysis of internal and external tissue damage ratio of the two opposite halves of the same fruit, according to the 1 to 5 severity scale; blue and red bars, pK7WG2::FaWRKY1/pK7WG2.0 and pK7WG2.0/pK7WG2.0 values, respectively. Data correspond to mean ± SD. Within each bars, means with different letters are significantly different by LSD test at $p < 0.05$. A ratio value of 1 indicate no differences between opposite halves of the same fruit.

4.4.5. Expression pattern of potential FaWRKY1 target genes in strawberry after *C. acutatum* inoculation

In order to uncover downstream FaWRKY1 defense responsive elements in strawberry, the expression of some strawberry orthologs of genes previously described in Arabidopsis as *WRKY75* target genes (Table S1) was evaluated after *C. acutatum* inoculation in non-agroinfiltrated fruit and in both *FaWRKY1* silenced and overexpressed fruit, at 4 and 5 days of agroinfiltration, where the highest levels of *FaWRKY1* transient silencing and overexpression were detected (Figures S2-S3). Only *FaCAT, FaWHY1, FaWHY2, FaJAZ9* and *FaJAZ5* genes responded positively to *C. acutatum* infection, and were significantly upregulated at 2 dpi (4dai) and/or 3 dpi (5 dai) in non agroinfiltrated fruit, being the expression of *FaJAZ4* significantly down-regulated at both times point (Figure 5). No significant changes in gene expression were observed for *FaJAZ1, FaJAZ8.1, FaJAZ10, FaJAZ12* and *FaICS1* neither at 2 dpi (4dai) nor 3 dpi (5dai) (data not shown). Also, no significant change in the expression of any of the strawberry orthologs was detected at 2 dpi and 3 dpi, when *FaWRKY1* was transiently silenced or overexpressed (data not shown).

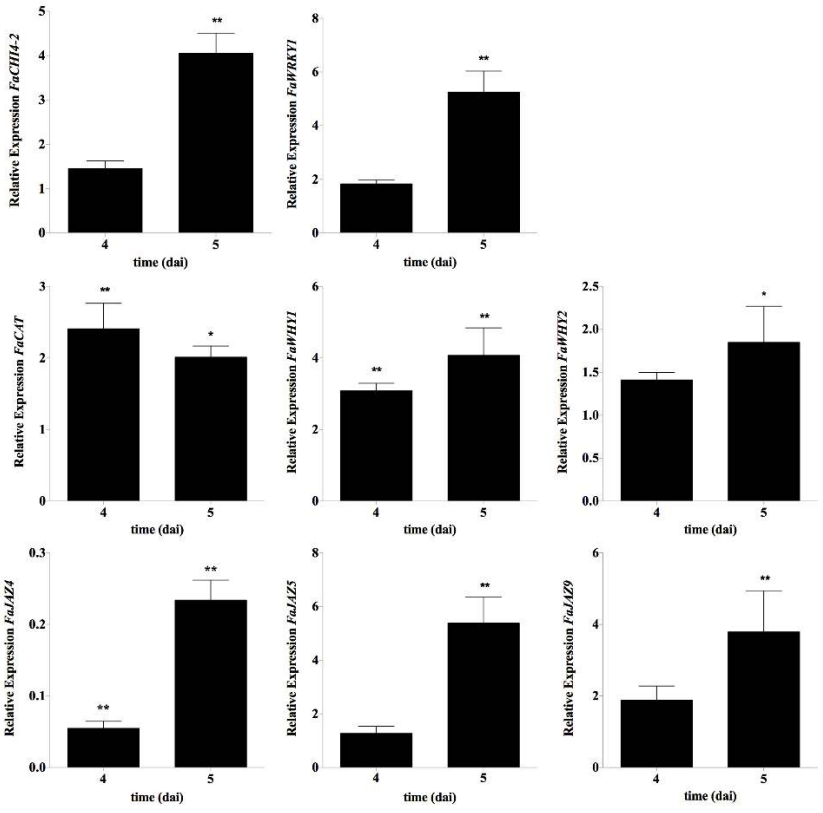


Figure 5. Relative expression patterns of potential *FaWRKY1*-responsive strawberry genes after *C. acutatum* inoculation. RT-qPCR analysis was accomplished in non-agroinfiltrated strawberry fruits. The time scale is shown as for agroinfiltration fruit samples (4 and 5 dai, corresponding to 2 and 3 days post inoculation with *C. acutatum*). *FaCHI4-2* and *FaWRKY1* were included as positive controls. Mean, standard error and significant differences found by Dunnett's test are represented (* p≤0.05, ** p≤0.01; n=3).

4.5. Discussion

To get insight into the role of *FaWRKY1* in the strawberry defense response to pathogens, we have accomplished the transient silencing and overexpression of this gene in strawberry fruits by *Agrobacterium* mediated transformation, which were later analyzed upon Colletotrichum acutatum inoculation. Transient expression by agroinfiltration has been applied to strawberry fruit as an efficient system to characterize genes associated with fruit development, physiology and metabolism (Carvalho et al., 2016). However, this methodology presents the limitation that *Agrobacterium* itself is an unusual plant pathogen, which can hamper the study of other plant-pathogen interactions (Guidarelli and Baraldi, 2015;Carvalho et al., 2016). Here, a modified experimental system where opposite halves of the same fruit are transiently and independently transformed has allowed the study of the effect of a transgene and its control in the same single fruit so that gene expression was only confined within the tissue area of the corresponding injected fruit half (Figures 1, 2 and Figure S1). This fact is particular interesting since to date, the process of infiltration has been carried out by a single injection of the *Agrobacterium* culture into the whole fruit, with a uniform expression of the transgene both close and distant from the injection sites (Hoffmann et al., 2006;Guidarelli et al., 2014). An advantage of using this technique is that the effect of a transgene produced in one half of the fruit can be normalized with respect to the control vector, in the opposite half (Figure S2, red scale) and variability is strongly reduced, as the physiological/developmental stages between opposite fruit halves are identical or closely identical.

4.5.1. The FaWRKY1 negatively regulates resistance to *C. acutatum* in strawberry fruit upon infection

FaWRKY1 gene expression was efficiently silenced after infiltration with *Agrobacterium* bearing the silencing cassette construct in both *C. acutatum* inoculated and non-inoculated fruit halves (Figure S2). As the level of silencing of *FaWRKY1* was relevant at 2 days after agroinfiltration (25% of gene silencing) in non-inoculated fruit, this time point was selected to inoculate *C. acutatum* for further fruit susceptibility assays, also bearing in mind that this pathogen takes around 24 hours to develop subcuticular intracellular hyphae (Guidarelli et al., 2011;Amil-Ruiz et al., 2016). In this way, we matched the initial phases of growth and development of the fungi with the time when a very high value of *FaWRKY1* silencing was detected in agroinfiltrated fruit.

The fruit susceptibility assays demonstrated that down-regulation of *FaWRKY1* gene in fruit enhances resistance to *C. acutatum* (Figure 3). Thus, tissue damage was visually reduced in fruit halves that were infiltrated with *Agrobacterium* bearing the silencing construct compared to that of the opposite control halves (Figure 3B). Also, most of the silenced fruit halves scored lower values of tissue damage (mainly 1 to 3) than control fruit halves, which grouped at higher values of severity (scores 3 to 5) (Figure 3C). Furthermore, when tissue damage values were normalized between opposite

halves, enhanced resistance was found statistically significant in the silenced fruit (Figure 3D). Taking together, these results evidence a role of FaWRKY1 in strawberry fruit as negative regulator of resistance to the pathogen *C. acutatum*. It is worthwhile to note that no clear difference in external surface damage was detected between silenced and control fruit. Although alternative explanations are plausible, it can presumably be due to a much lower transformation event produced in surface fruit cells than in internal tissues cells as a consequence of the agroinfiltration procedure.

4.5.2. Ectopic expression of *FaWRKY1* in strawberry fruit does not increase susceptibility to *C. acutatum*

The transient overexpression of *FaWRKY1* gene in response to *C. acutatum* did not result in any change in fruit susceptibility to this pathogen (Figure 4) even though a clear and significant accumulation of *FaWRKY1* transcripts was detected at 4 and 5 days after agroinfiltration (Figure S3). Thus, neither external nor internal clear distinguishing morphological differences were visually observed between fruit halves infiltrated with *Agrobacterium* bearing the overexpression cassette and the corresponding opposite fruit halves bearing the empty vector, as control (Figure 4A and B). The distribution pattern of fruit halves according to their tissue damage was also similar for both overexpressed and control fruit halves (Figure 4C) and no significant differences were found when tissue damage values were normalized between the two opposite halves of the same fruit (Figure 4D).

Although increase of fruit susceptibility after *FaWRKY1* overexpression could be expected in fruit cells, however, the presence of high levels of this transcription factor acting as a gene repressor may not necessarily end in increased susceptibility to this pathogen. In normal uninfected conditions, a basal *FaWRKY1* expression in plant cells can be enough to have the complete set of specific FaWRKY1-responsive defense genes repressed in order to optimize plant fitness (Huot et al., 2014). Thus, high levels of FaWRKY1 proteins present in fruit cells after *FaWRKY1* ectopic expression would not add more repression on these *FaWRKY1*-related genes but this repression would remain unaltered. Previous studies in *Arabidopsis* had revealed a repressor role of FaWRKY1 on defense related genes when this *FaWRKY1* gene was ectopically overexpressed in wild type plants (Encinas-Villarejo et al., 2009). However, in that study, the overexpression of strawberry *FaWRKY1* gene in *Arabidopsis Atwrky75* mutant, restored the susceptible phenotype to wild-type and even increased resistance of the mutant to avirulent strains of *P. syringae*, (Encinas-Villarejo et al., 2009). A similar pattern of pathogen resistant has also been described by Guo et al. (2017) in *Arabidopsis* WRKY75ox (over-expressing) lines, which show an elevated SA content and enhanced resistance to Pst DC3000, compared with both, wild type Col-0 and WRKY75RNAi plants.

What are the molecular events underlying a different plant defense response after the ectopic expression of *FaWRKY1* in strawberry and Arabidopsis remain to be further

elucidated. Multiple function variability and positive or negative regulator roles have been described for AtWRKY75 and its homologs in other species. Interestingly, high levels of gene expression were detected for the *WRKY75*-like orthologous gene in a peach resistant cultivar to *Xanthomonas arboricola* compared with a more susceptible cultivar (Gervasi et al., 2018). Also, overexpression of *VvWRKY1* (*AtWRKY75*-like) in grapevines enhanced resistance to biotrophic *Plasmopara viticola*, the causal agent of downy mildew, through induction of JA-pathway related genes (Marchive et al., 2013). Contrarily, and accordingly to our results in the present report, the silencing of *GbWRKY1* in cotton enhanced plant resistance to hemibiotrophic fungal *V. dahliae* and *B. cinerea* (Marchive et al., 2013;Li et al., 2014).

Taken together all these results, one might speculate on WRKY75-like TFs acting as a positive regulator of resistance against both bacterial and biotrophic fungi pathogens, whereas it negatively regulates defense responses against fungi with a necrotrophic phase or lifestyles. However, it is also worthwhile to note that the positive regulator role of FaWRKY1 in pathogen resistance observed in *Arabidopsis* was detected in tissue plant other than fruit. It is known that hormones modulate plant immunity, with SA and JA as major players. However, ethylene, abscisic acid, gibberellins, auxins, cytokinins, brassinosteroids and nitric oxide, also have pivotal roles in the regulation of the plant immune signaling network (Pieterse et al., 2012). Moreover, interplay between phytohormones is required for development, maturation, and ripening of fruit (McAtee et al., 2013) and thus, strawberry fruit tissue undertakes substantial changes in the hormonal balance over growing and ripening time (Medina-Puche et al., 2016), which could affect main defense pathways differently from other tissues.

Together, our results in strawberry indicate that the overexpression of *FaWRKY1* in fruit does not substantially affect fruit susceptibility to *C. acutatum*, whereas strong evidence is also provided that the silencing of *FaWRKY1* in strawberry fruit enhances resistance to *C. acutatum*. Also, these results evidence the complexity and multiple layers of control that FaWRKY1 can exhibit and highlight differences in the defense response strategies activated either by AtWRKY75 or FaWRKY1 proteins according to different plant species, plant tissue and/or different style of life deployed by *P. syringae* or *C. acutatum* pathogens, respectively.

4.5.3. Downstream defense responsive elements and underlying mechanisms of *WRKY75*-like genes and *FaWRKY1* in strawberry

As mentioned before, and accordingly to our results in strawberry, the silencing of *GbWRKY1* in cotton has been reported to enhance plant resistance to hemibiotrophic fungal *V. dahliae* and *B. cinerea* (Marchive et al., 2013;Li et al., 2014). Interestingly, GbWRKY1 acted by promoting transcription of *JAZ1* homologs. JA ZIM-domain (JAZ) family proteins are repressors interacting with several transcriptional factors involved in the regulation of early JA-responsive genes (Chini et al., 2007;Fernandez-Calvo et al., 2011;Major et al., 2017). Thus, at low JA-Ile levels, JAZ protein negatively regulates JA

signaling pathway, including many TFs which positively regulate JA-responsive genes (Gimenez-Ibanez et al., 2016), whereas in presence of active form of JA (JA-Ile), JAZ proteins are targeted by the SCFCOI1 complex, and are subsequently degraded by the 26S proteasome (Chini et al., 2007;Thines et al., 2007;Kazan and Manners, 2012).

On the other hand, WRKY75 positively regulates plant resistance to *P. syringae* in *Arabidopsis*, but is also upregulated during leaf senescence, a complicated process influenced by a large number of genes, environment, stresses and endogenous levels of phytohormones (Guo et al., 2017;Li et al., 2017). This senescence process is regulated by WRKY75, which promotes the SA biosynthesis, directly activating the transcription of *Isochorismate synthase1* (*ICS1* or *SID2*), and H_2O_2 accumulation (Guo et al., 2017;Li et al., 2017). Interestingly, *FaWRKY1* is also up-regulated in over ripen strawberry fruit (Encinas-Villarejo et al., 2009). Also, AtWRKY75 is able to suppress catalase activity by directly repressing *CAT2* transcription, which directly contributes to increase the production of ROS (Guo et al., 2017;Li et al., 2017). High levels of ROS strongly correlate with the induction and maintenance of the cell senescence process and the hypersensitive response (HR), as also described elsewhere (Jajic et al., 2015). Furthermore, positive regulation of the defense response has been found recently involving the cassava WRKY75 homolog (MeWRKY75) interacting with Whirly (WHY) TF (Liu et al., 2018). MeWRKY75 positively regulates disease resistance to cassava bacterial blight first activating the expression of *MeWHY3* gene through directly binding to the W-box of its promoter region, and then promoting its physical interaction with MeWHYs. The physical interaction between MeWHYs and MeWRKY75 contributes to the activation of defense-related genes and improves resistance against the biotrophic *Xanthomonas axonopodis* (Liu et al., 2018). Interestingly, AtWHYs are also involved in modulating leaf senescence (Miao et al., 2013;Ren et al., 2017).

Based on the results described above for AtWRKY75 and its homologs in plants, a tentative emerging proposal for the regulatory mechanism played by *AtWRKY75*-like genes in plant defense response in green tissues is described in Figure 6. It is worthwhile to note that this proposal only summarizes previous results described so far for *AtWRKY75* and *WRKY75*-like genes in other plant species. According to this tentative model and being FaWRKY1 a WRKY75-like transcription factor, it is not unreasonable to think that FaWRKY1 also might be involved in the up-regulation of certain *JAZ* genes such as *FaJAZ5* and *FaJAZ9* in strawberry. Therefore, when FaWRKY1 is present, a subsequent limited repression of genes involved in the early response of the JA-mediated pathway might be produced. Notably, under *C. acutatum* fruit infection the expression of *FaWRKY1* is upregulated (Figures 2 and 5, and (Encinas-Villarejo et al., 2009) and upon this pathogen interaction, a partial activation of JA-defensive pathway has been previously described in strawberry (Amil-Ruiz et al., 2016). Also in strawberry, a correlation between JA-Ile levels and expression pattern of some *JAZ* encoding genes (*FaJAZ1/8.1*) during fruit development, and JA-treated fruit has been reported (Garrido-Bigotes et al., 2017;Garrido-Bigotes et al., 2018) opening the possibility that increase in JA-Ile content by pathogen attack could up-regulate *JAZ* expression irrespectively of the WRKY75-like mediated control. In addition, also

similarly to AtWRKY75, FaWRKY1 might stimulate the biosynthesis of SA through upregulation of the *FaICS1* ortholog, which in turn will promote the activation of the SA-mediated pathway and the production of ROS. Notably, increases in SA and JA content has been described in strawberry after *C. acutatum* infection (Amil-Ruiz et al., 2016), and the ectopic expression of *FaWRKY1* in *Arabidopsis* wild-type and WRKY75At22 mutant genetic backgrounds promoted the production of high levels of H_2O_2 after being challenged with *P. syringae* (Encinas-Villarejo et al., 2009).

In an attempt to address whether FaWRKY1 suits to this tentative model in strawberry fruit and expand our understanding of potential molecular players within this complex regulatory network, we have monitored the molecular signature of *ICS1*, *CAT2*, *JAZ* and *WHY* family orthologs in strawberry and in both *FaWRKY1* silenced and overexpressed fruit. Thus, we have identified one *FaICS1*, one *FaCAT* and two *FaWHY* genes in strawberry and twelve ortholog members of the JAZ protein family previously described and characterized in the diploid woodland strawberry *F. vesca* (Garrido-Bigotes et al., 2018). Only *FaCAT, FaWHY1, FaWHY2, FaJAZ9* and *FaJAZ5* genes responded positively to *C. acutatum* infection in non agroinfiltrated fruit (Figure 5). Moreover, we could not detect significant changes in gene expression for any of the strawberry tested genes in fruit in which *FaWRKY1* was transiently silenced or overexpressed. Therefore, in our experimental conditions, no clear positive or negative correlation can be inferred between the FaWRKY1 and the strawberry *FaICS1, FaCAT, FaWHY* or *FaJAZ* genes here analyzed and this matter remains to be further elucidated.

The non-climacteric strawberry fruit growing and ripening process involve major physiological, chemical and programmed hormonal changes impacting the defense mechanisms (Amil-Ruiz et al., 2011), differing from other climacteric species and plant tissues. Remarkably, an ABA increase in the red fruit stage takes place (Symons et al., 2012). It has been shown that ABA mediates defense responses positive or negatively depending on the pathogen life style and tissue infected. Studies in Arabidopsis have shown that, after pathogen penetration, ABA antagonizes the SA-dependent defenses effective against biotrophic and hemibiotrophic pathogens. On the other hand, ABA promotes the MYC branch of the JA pathway while suppresses the ERF1/ORA59 branch, compromising the resistance to necrotrophs (Ton et al., 2009;Pieterse et al., 2012). Consequently, due to the fruit ripening process the expression pattern of defense related genes against *C. acutatum* and other pathogen infections could be modulated differently than in non-ripe fruit and vegetative tissues, by an altered balance between the antagonistic SA and JA responses and thus, facilitating the infection by necrotrophs (Alkan and Fortes, 2015). The ABA-mediated enhanced susceptibility of tomato fruit to *B. cinerea* has been described previously (Blanco-Ulate et al., 2013). Interestingly, the ABA pathway is also activated by *C. gloeosporioides* (Alkan et al., 2015). Besides, SA pathway activation by *C. coccodes* in ripe tomato fruit induced cell death, as well as it suppressed the JA pathway, playing an important role in the necrotrophic development of this fungus (Alkan et al., 2012). Within this intricate network of pathogen effectors, antagonistic defense pathways activation and hormones crosstalk, FaWRKY1 seems to have a role in regulating the SA and JA defense

pathways balance in the strawberry fruit, albeit further research is needed to fully understand the mechanisms and the players involved.

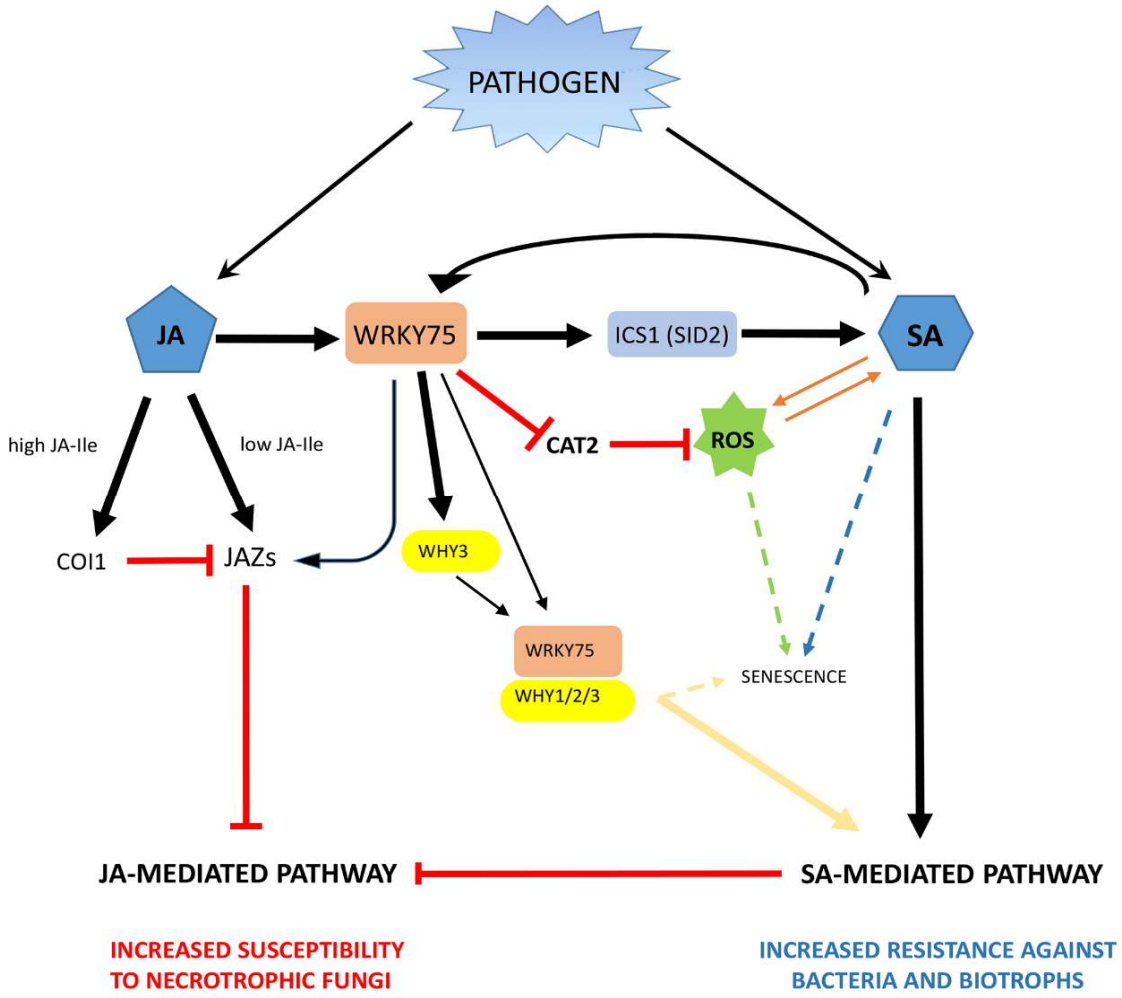


Figure 6. Schematic of a tentative model for downstream regulatory networks dictated by AtWRKY75-like genes in plants. The model is based on the results described for AtWRKY75 and its homologs in green tissue from species other than strawberry. Thus, the cotton GbWRKY1 (AtWRKY75-like) can activate JAZ1 expression, which interferes with the JA-mediated defense pathway and negatively regulate plant resistance to the pathogens B. cinerea and V. dahliae (Li et al., 2014). In addition, AtWRKY75 directly bind to the *Isochorismate synthase1* (*SID2*) promoter to positively regulate its transcription and stimulate the biosynthesis of SA, which in turn seems to promote the activation of SA-mediated pathway and the generation of reactive oxygen species (ROS) and positively regulate plant resistance to P. syringae (Guo et al., 2017). Also, AtWRKY75 is able to suppress catalase activity by directly repressing *CAT2* transcription, which directly contributes to increase the production of ROS (Guo et al., 2017). Other interactions of WRKY75 homologs, includes WHIRLY factors, as described for MeWRKY75 (Liu et al., 2018). Thus, MeWRKY75 is able to activate MeWHY3 transcription and also physically interact with WHY factors to form a protein complex and mediate disease resistance to cassava bacterial blight infection (Liu et al., 2018). Interestingly, WHY1 has also been associated to senescence processes in *Arabidopsis* (Ren et al. 2017). As a result, the SA-mediated defense pathway is promoted, hence the resistance to bacteria and biotrophic fungi might increase, and the JA-mediated pathway is antagonized, increasing the susceptibility to necrotrophic pathogens. Solid arrows denote direct positive regulation of genes or pathways. "T" lines mean negative regulation. Dashed lines evidence positive correlation between plant events.

4.6. Conclusions

In summary, a functional characterization of the *FaWRKY1* gene has been accomplished in strawberry fruit. We provide evidences that suggest the relevance between FaWRKY1 and strawberry fruit disease resistance against *C. acutatum*. Thus, FaWRKY1 act as a negative regulator of strawberry fruit resistance to *C. acutatum*. The FaWRKY1 responsive elements and molecular mechanisms involved in the defense response to *C. acutatum* remain elusive and further studies are still needed to unravel the intricate and complex regulatory network of FaWRKY1 in strawberry.

4.7. References

Aharoni, A., Keizer, L.C., Van Den Broeck, H.C., Blanco-Portales, R., Munoz-Blanco, J., Bois, G., et al. (2002). Novel insight into vascular, stress, and auxin-dependent and -independent gene expression programs in strawberry, a non-climacteric fruit. *Plant Physiol* 129**,** 1019-1031.doi: 10.1104/pp.003558

Alkan, N., Fluhr, R., and Prusky, D. (2012). Ammonium secretion during Colletotrichum coccodes infection modulates salicylic and jasmonic acid pathways of ripe and unripe tomato fruit. *Mol Plant Microbe Interact* 25**,** 85-96.doi: 10.1094/MPMI-01-11-0020

Alkan, N., and Fortes, A.M. (2015). Insights into molecular and metabolic events associated with fruit response to post-harvest fungal pathogens. *Frontiers in Plant Science* 6**,** 889

Alkan, N., Friedlander, G., Ment, D., Prusky, D., and Fluhr, R. (2015). Simultaneous transcriptome analysis of Colletotrichum gloeosporioides and tomato fruit pathosystem reveals novel fungal pathogenicity and fruit defense strategies. *New Phytol* 205**,** 801-815.doi: 10.1111/nph.13087

Amil-Ruiz, F., Blanco-Portales, R., Munoz-Blanco, J., and Caballero, J.L. (2011). The strawberry plant defense mechanism: a molecular review. *Plant Cell Physiol* 52**,** 1873-1903.doi: 10.1093/pcp/pcr136

Amil-Ruiz, F., Garrido-Gala, J., Blanco-Portales, R., Folta, K.M., Munoz-Blanco, J., and Caballero, J.L. (2013). Identification and validation of reference genes for transcript normalization in strawberry (Fragaria x ananassa) defense responses. *PLoS One* 8**,** e70603.doi: 10.1371/journal.pone.0070603

Amil-Ruiz, F., Garrido-Gala, J., Gadea, J., Blanco-Portales, R., Muñoz-Mérida, A., Trelles, O., et al. (2016). Partial Activation of SA- and JA-Defensive Pathways in Strawberry upon Colletotrichum acutatum Interaction. *Frontiers in Plant Science* 7**,** 1036

Blanco-Ulate, B., Vincenti, E., Powell, A.L., and Cantu, D. (2013). Tomato transcriptome and mutant analyses suggest a role for plant stress hormones in the interaction between fruit and Botrytis cinerea. *Front Plant Sci* 4**,** 142.doi: 10.3389/fpls.2013.00142

Carvalho, R.F., Carvalho, S.D., O'grady, K., and Folta, K.M. (2016). Agroinfiltration of Strawberry Fruit — A Powerful Transient Expression System for Gene Validation. *Current Plant Biology* 6**,** 19-37.doi: 10.1016/j.cpb.2016.09.002

Casado-Díaz, A., Encinas-Villarejo, S., Santos, B.D.L., Schilirò, E., Yubero-Serrano, E.-M., Amil-Ruíz, F., et al. (2006). Analysis of strawberry genes differentially expressed in response to Colletotrichum infection. *Physiologia Plantarum* 128**,** 633-650.doi: 10.1111/j.1399-3054.2006.00798.x

Chini, A., Fonseca, S., Fernandez, G., Adie, B., Chico, J.M., Lorenzo, O., et al. (2007). The JAZ family of repressors is the missing link in jasmonate signalling. *Nature* 448**,** 666-671.doi: 10.1038/nature06006

Curry, K.J., Abril, M., Avant, J.B., and Smith, B.J. (2002). Strawberry Anthracnose: Histopathology of Colletotrichum acutatum and C. fragariae. *Phytopathology* 92**,** 1055-1063.doi: 10.1094/PHYTO.2002.92.10.1055

Desveaux, D., Marechal, A., and Brisson, N. (2005). Whirly transcription factors: defense gene regulation and beyond. *Trends Plant Sci* 10**,** 95-102.doi: 10.1016/j.tplants.2004.12.008

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. *Trends Plant Sci* 15**,** 573-581.doi: 10.1016/j.tplants.2010.06.005

Encinas-Villarejo, S., Maldonado, A.M., Amil-Ruiz, F., De Los Santos, B., Romero, F., Pliego-Alfaro, F., et al. (2009). Evidence for a positive regulatory role of strawberry (Fragaria×ananassa) Fa WRKY1 and Arabidopsis At WRKY75 proteins in resistance. *Journal of Experimental Botany* 60**,** 3043-3065.doi: 10.1093/jxb/erp152

Eulgem, T., and Somssich, I.E. (2007). Networks of WRKY transcription factors in defense signaling. *Curr Opin Plant Biol* 10**,** 366-371.doi: 10.1016/j.pbi.2007.04.020

Fernandez-Calvo, P., Chini, A., Fernandez-Barbero, G., Chico, J.M., Gimenez-Ibanez, S., Geerinck, J., et al. (2011). The Arabidopsis bHLH transcription factors MYC3 and MYC4 are targets of JAZ repressors and act additively with MYC2 in the activation of jasmonate responses. *Plant Cell* 23**,** 701-715.doi: 10.1105/tpc.110.080788

Fu, Z.Q., and Dong, X. (2013). Systemic acquired resistance: turning local infection into global defense. *Annu Rev Plant Biol* 64**,** 839-863.doi: 10.1146/annurev-arplant-042811-105606

Gao, Q.M., Venugopal, S., Navarre, D., and Kachroo, A. (2011). Low oleic acid-derived repression of jasmonic acid-inducible defense responses requires the WRKY50 and WRKY51 proteins. *Plant Physiol* 155**,** 464-476.doi: 10.1104/pp.110.166876

Garrido-Bigotes, A., Figueroa, N.E., Figueroa, P.M., and Figueroa, C.R. (2018). Jasmonate signalling pathway in strawberry: Genome-wide identification, molecular characterization and expression of JAZs and MYCs during fruit development and ripening. *PLoS One* 13**,** e0197118.doi: 10.1371/journal.pone.0197118

Garrido-Bigotes, A., Figueroa, P.M., and Figueroa, C.R. (2017). Jasmonate Metabolism and Its Relationship with Abscisic Acid During Strawberry Fruit Development and Ripening. *Journal of Plant Growth Regulation* 37**,** 101-113.doi: 10.1007/s00344-017-9710-x

Gervasi, F., Ferrante, P., Dettori, M.T., Scortichini, M., and Verde, I. (2018). Transcriptome reprogramming of resistant and susceptible peach genotypes during Xanthomonas arboricola pv. pruni early leaf infection. *PLoS One* 13**,** e0196590.doi: 10.1371/journal.pone.0196590

Giampieri, F., Alvarez-Suarez, J.M., Cordero, M.D., Gasparrini, M., Forbes-Hernandez, T.Y., Afrin, S., et al. (2017). Strawberry consumption improves aging-associated impairments, mitochondrial biogenesis and functionality through the AMP-activated protein kinase signaling cascade. *Food Chem* 234**,** 464-471.doi: 10.1016/j.foodchem.2017.05.017

Gimenez-Ibanez, S., Chini, A., and Solano, R. (2016). How Microbes Twist Jasmonate Signaling around Their Little Fingers. *Plants (Basel)* 5.doi: 10.3390/plants5010009

Guidarelli, M., and Baraldi, E. (2015). Transient transformation meets gene function discovery: the strawberry fruit case. *Front Plant Sci* 6**,** 444.doi: 10.3389/fpls.2015.00444

Guidarelli, M., Carbone, F., Mourgues, F., Perrotta, G., Rosati, C., Bertolini, P., et al. (2011). Colletotrichum acutatum interactions with unripe and ripe strawberry fruits and differential responses at histological and transcriptional levels. *Plant Pathology* 60**,** 685-697.doi: 10.1111/j.1365-3059.2010.02423.x

Guidarelli, M., Zoli, L., Orlandini, A., Bertolini, P., and Baraldi, E. (2014). The mannose-binding lectin gene FaMBL1 is involved in the resistance of unripe strawberry fruits to Colletotrichum acutatum. *Mol Plant Pathol* 15**,** 832-840.doi: 10.1111/mpp.12143

Guo, P., Li, Z., Huang, P., Li, B., Fang, S., Chu, J., et al. (2017). A Tripartite Amplification Loop Involving the Transcription Factor WRKY75, Salicylic Acid, and Reactive Oxygen Species Accelerates Leaf Senescence. *Plant Cell* 29**,** 2854-2870.doi: 10.1105/tpc.17.00438

Hoffmann, T., Kalinowski, G., and Schwab, W. (2006). RNAi-induced silencing of gene expression in strawberry fruit (Fragaria x ananassa) by agroinfiltration: a rapid assay for gene function analysis. *Plant J* 48**,** 818-826.doi: 10.1111/j.1365-313X.2006.02913.x

Holsters, M., De Waele, D., Depicker, A., Messens, E., Van Montagu, M., and Schell, J. (1978). Transfection and transformation of Agrobacterium tumefaciens. *Mol Gen Genet* 163**,** 181-187.doi: 10.1007/BF00267408

Horowitz, S., Freeman, S., and Sharon, A. (2002). Use of Green Fluorescent Protein-Transgenic Strains to Study Pathogenic and Nonpathogenic Lifestyles in Colletotrichum acutatum. *Phytopathology* 92**,** 743-749.doi: 10.1094/PHYTO.2002.92.7.743

Huot, B., Yao, J., Montgomery, B.L., and He, S.Y. (2014). Growth-defense tradeoffs in plants: a balancing act to optimize fitness. *Mol Plant* 7**,** 1267-1287.doi: 10.1093/mp/ssu049

Ishihama, N., and Yoshioka, H. (2012). Post-translational regulation of WRKY transcription factors in plant immunity. *Curr Opin Plant Biol* 15**,** 431-437.doi: 10.1016/j.pbi.2012.02.003

Jajic, I., Sarna, T., and Strzalka, K. (2015). Senescence, Stress, and Reactive Oxygen Species. *Plants (Basel)* 4**,** 393-411.doi: 10.3390/plants4030393

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO journal* 6**,** 3901-3907

Jeger, M.J., and Bailey, J.A. (1992). "Colletotrichum: biology, pathology, and control/Edited for the British Society for Plant Pathology by JA Bailey and MJ Jeger". British Society for Plant Pathology.).

Jin Choi, H., Seuk Bae, Y., Soo Lee, J., Hea Park, M., and Gang Kim, J. (2016). Effects of Carbon Dioxide Treatment and Modified Atmosphere Packaging on the Quality of Long Distance Transporting “Maehyang” Strawberry. *Agricultural Sciences* 07**,** 813-821.doi: 10.4236/as.2016.712074

Kazan, K., and Manners, J.M. (2012). JAZ repressors and the orchestration of phytohormone crosstalk. *Trends Plant Sci* 17**,** 22-31.doi: 10.1016/j.tplants.2011.10.006

Lai, Z., Vinod, K., Zheng, Z., Fan, B., and Chen, Z. (2008). Roles of Arabidopsis WRKY3 and WRKY4 transcription factors in plant responses to pathogens. *BMC Plant Biol* 8**,** 68.doi: 10.1186/1471-2229-8-68

Lazo, G.R., Stein, P.A., and Ludwig, R.A. (1991). A DNA transformation-competent Arabidopsis genomic library in Agrobacterium. *Biotechnology (N Y)* 9**,** 963-967.doi: 10.1038/nbt1091-963

Li, C., He, X., Luo, X., Xu, L., Liu, L., Min, L., et al. (2014). Cotton WRKY1 mediates the plant defense-to-development transition during infection of cotton by Verticillium dahliae by activating JASMONATE ZIM-DOMAIN1 expression. *Plant Physiol* 166**,** 2179-2194.doi: 10.1104/pp.114.246694

Li, J., Brader, G., Kariola, T., and Palva, E.T. (2006). WRKY70 modulates the selection of signaling pathways in plant defense. *Plant J* 46**,** 477-491.doi: 10.1111/j.1365-313X.2006.02712.x

Li, J., Brader, G., and Palva, E.T. (2004). The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. *Plant Cell* 16**,** 319-331.doi: 10.1105/tpc.016980

Li, Z., Zhao, Y., Liu, X., Jiang, Z., Peng, J., Jin, J., et al. (2017). Construction of the Leaf Senescence Database and Functional Assessment of Senescence-Associated Genes. *Methods Mol Biol* 1533**,** 315-333.doi: 10.1007/978-1-4939-6658-5_19

Liu, W., Yan, Y., Zeng, H., Li, X., Wei, Y., Liu, G., et al. (2018). Functional characterization of WHY-WRKY75 transcriptional module in plant response to cassava bacterial blight. *Tree Physiol* 38**,** 1502-1512.doi: 10.1093/treephys/tpy053

Liu, X., Song, Y., Xing, F., Wang, N., Wen, F., and Zhu, C. (2016). GhWRKY25, a group I WRKY gene from cotton, confers differential tolerance to abiotic and biotic stresses in transgenic Nicotiana benthamiana. *Protoplasma* 253**,** 1265-1281.doi: 10.1007/s00709-015-0885-3

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25**,** 402-408.doi: 10.1006/meth.2001.1262

Maas, J.L. (1998). Compendium of Strawberry Diseases. *The American Phytopathological Society.St. Paul, Minnesota, USA* Second Edition. pp: 44-57

Major, I.T., Yoshida, Y., Campos, M.L., Kapali, G., Xin, X.F., Sugimoto, K., et al. (2017). Regulation of growth-defense balance by the JASMONATE ZIM-DOMAIN (JAZ)-MYC transcriptional module. *New Phytol* 215**,** 1533-1547.doi: 10.1111/nph.14638

Marchive, C., Leon, C., Kappel, C., Coutos-Thevenot, P., Corio-Costet, M.F., Delrot, S., et al. (2013). Over-expression of VvWRKY1 in grapevines induces expression of jasmonic acid pathway-related genes and confers higher tolerance to the downy mildew. *PLoS One* 8**,** e54185.doi: 10.1371/journal.pone.0054185

Mcatee, P., Karim, S., Schaffer, R., and David, K. (2013). A dynamic interplay between phytohormones is required for fruit development, maturation, and ripening. *Front Plant Sci* 4**,** 79.doi: 10.3389/fpls.2013.00079

Medina-Puche, L., Blanco-Portales, R., Molina-Hidalgo, F.J., Cumplido-Laso, G., Garcia-Caparros, N., Moyano-Canete, E., et al. (2016). Extensive transcriptomic studies on the roles played by abscisic acid and auxins in the development and ripening of strawberry fruits. *Funct Integr Genomics* 16**,** 671-692.doi: 10.1007/s10142-016-0510-3

Miao, Y., Jiang, J., Ren, Y., and Zhao, Z. (2013). The single-stranded DNA-binding protein WHIRLY1 represses WRKY53 expression and delays leaf senescence in a developmental stage-dependent manner in Arabidopsis. *Plant Physiol* 163**,** 746-756.doi: 10.1104/pp.113.223412

Neu, E., Domes, H.S., Menz, I., Kaufmann, H., Linde, M., and Debener, T. (2019). Interaction of roses with a biotrophic and a hemibiotrophic leaf pathogen leads to differences in defense transcriptome activation. *Plant Mol Biol* 99**,** 299-316.doi: 10.1007/s11103-018-00818-2

Pandey, S.P., and Somssich, I.E. (2009). The role of WRKY transcription factors in plant immunity. *Plant Physiol* 150**,** 1648-1655.doi: 10.1104/pp.109.138990

Peres, N.A., Timmer, L.W., Adaskaveg, J.E., and Correll, J.C. (2005). Lifestyles of Colletotrichum acutatum. *Plant Dis* 89**,** 784-796.doi: 10.1094/PD-89-0784

Pieterse, C.M., Van Der Does, D., Zamioudis, C., Leon-Reyes, A., and Van Wees, S.C. (2012). Hormonal modulation of plant immunity. *Annu Rev Cell Dev Biol* 28**,** 489-521.doi: 10.1146/annurev-cellbio-092910-154055

Ren, Y., Li, Y., Jiang, Y., Wu, B., and Miao, Y. (2017). Phosphorylation of WHIRLY1 by CIPK14 Shifts Its Localization and Dual Functions in Arabidopsis. *Mol Plant* 10**,** 749-763.doi: 10.1016/j.molp.2017.03.011

Robatzek, S., and Somssich, I.E. (2002). Targets of AtWRKY6 regulation during plant senescence and pathogen defense. *Genes Dev* 16**,** 1139-1149.doi: 10.1101/gad.222702

Robert-Seilaniantz, A., Grant, M., and Jones, J.D. (2011). Hormone crosstalk in plant disease and defense: more than just jasmonate-salicylate antagonism. *Annu Rev Phytopathol* 49**,** 317-343.doi: 10.1146/annurev-phyto-073009-114447

Ruijter, J.M., Ramakers, C., Hoogaars, W.M., Karlen, Y., Bakker, O., Van Den Hoff, M.J., et al. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res* 37**,** e45.doi: 10.1093/nar/gkp045

Rushton, P.J., Somssich, I.E., Ringler, P., and Shen, Q.J. (2010). WRKY transcription factors. *Trends Plant Sci* 15**,** 247-258.doi: 10.1016/j.tplants.2010.02.006

Seo, E., Choi, D., and Choi (2015). Functional studies of transcription factors involved in plant defenses in the genomics era. *Brief Funct Genomics* 14**,** 260-267.doi: 10.1093/bfgp/elv011

Simpson, D.W. (2015). Resistance toBotrytis cinereain pistillate genotypes of the cultivated strawberryFragaria ananassa. *Journal of Horticultural Science* 66**,** 719-723.doi: 10.1080/00221589.1991.11516203

Spolaore, S., Trainotti, L., and Casadoro, G. (2001). A simple protocol for transient gene expression in ripe fleshy fruit mediated by Agrobacterium. *J Exp Bot* 52**,** 845-850.doi: 10.1093/jexbot/52.357.845

Sreenivasaprasad, S., and Talhinhas, P. (2005). Genotypic and phenotypic diversity in Colletotrichum acutatum, a cosmopolitan pathogen causing anthracnose on a wide range of hosts. *Mol Plant Pathol* 6**,** 361-378.doi: 10.1111/j.1364-3703.2005.00291.x

Symons, G.M., Chua, Y.J., Ross, J.J., Quittenden, L.J., Davies, N.W., and Reid, J.B. (2012). Hormonal changes during non-climacteric ripening in strawberry. *J Exp Bot* 63**,** 4741-4750.doi: 10.1093/jxb/ers147

Thines, B., Katsir, L., Melotto, M., Niu, Y., Mandaokar, A., Liu, G., et al. (2007). JAZ repressor proteins are targets of the SCF(COI1) complex during jasmonate signalling. *Nature* 448**,** 661-665.doi: 10.1038/nature05960

Ton, J., Flors, V., and Mauch-Mani, B. (2009). The multifaceted role of ABA in disease resistance. *Trends Plant Sci* 14**,** 310-317.doi: 10.1016/j.tplants.2009.03.006

Wang, D., Amornsiripanitch, N., and Dong, X. (2006). A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants. *PLoS Pathog* 2**,** e123.doi: 10.1371/journal.ppat.0020123

Wang, X., Guo, R., Tu, M., Wang, D., Guo, C., Wan, R., et al. (2017). Ectopic Expression of the Wild Grape WRKY Transcription Factor VqWRKY52 in Arabidopsis thaliana Enhances Resistance to the Biotrophic Pathogen Powdery Mildew But Not to the Necrotrophic Pathogen Botrytis cinerea. *Front Plant Sci* 8**,** 97.doi: 10.3389/fpls.2017.00097

Wei, W., Hu, Y., Han, Y.T., Zhang, K., Zhao, F.L., and Feng, J.Y. (2016). The WRKY transcription factors in the diploid woodland strawberry Fragaria vesca: Identification and expression analysis under biotic and abiotic stresses. *Plant Physiol Biochem* 105**,** 129-144.doi: 10.1016/j.plaphy.2016.04.014

Yan, C., and Xie, D. (2015). Jasmonate in plant defence: sentinel or double agent? *Plant Biotechnol J* 13**,** 1233-1240.doi: 10.1111/pbi.12417

Zheng, Z., Qamar, S.A., Chen, Z., and Mengiste, T. (2006). Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant J* 48**,** 592-605.doi: 10.1111/j.1365-313X.2006.02901.x

Zhou, M., and Memelink, J. (2016). Jasmonate-responsive transcription factors regulating plant secondary metabolism. *Biotechnol Adv* 34**,** 441-449.doi: 10.1016/j.biotechadv.2016.02.004

Zhu, L., Ni, W., Liu, S., Cai, B., Xing, H., and Wang, S. (2017). Transcriptomics Analysis of Apple Leaves in Response to Alternaria alternata Apple Pathotype Infection. *Front Plant Sci* 8**,** 22.doi: 10.3389/fpls.2017.00022

Capítulo 5:

Conclusiones

Conclusiones.

1. Se han identificado 64 y 257 genes que codifican factores de transcripción WRKY en los genomas de *Fragaria vesca* y *Fragaria x ananassa*, respectivamente, incluyendo por primera vez varios genes parálogos y, en el caso del octoploide, también sus genes homoeólogos. Además de conservar una elevada homología en sus secuencias, ambos grupos de genes son, en su mayoría, sinténicos y colineares. Además, están sometidos fundamentalmente a una presión de selección negativa (purificante) que favorece la conservación de sus funciones. La formación de los híbridos octoploides ha originado nuevas duplicaciones génicas que formaron los parálogos en *Fragaria x ananassa* no presentes en la especie diploide.
2. Los diferentes patrones de expresión y la expresión diferencial de los genes WRKY en varios tejidos indican que cumplen diferentes funciones en la regulación de la fisiología y el desarrollo de la planta de fresa y, concretamente, en la maduración del fruto y en los mecanismos de defensa frente al hongo hemibiotrofo *Colletotrichum*. La expresión de los parálogos y homoeologos sugiere que actúan de forma aditiva en algunos casos, mientras que en otros, las diferencias en su expresión apuntan a fenómenos de neo- o subfuncionalización.
3. Existen 25 y 96 genes que codifican proteínas VQ en los genomas de *Fragaria vesca* y *Fragaria x ananassa*, respectivamente, incluyendo varios genes parálogos y, en el caso del octoploide, también homoeólogos. Además de conservar una elevada homología en sus secuencias, ambos grupos de genes son, en su mayoría, sinténicos y colineares. Además, están sometidos fundamentalmente a una presión de selección negativa (purificante) que favorece la conservación de sus funciones.
4. Se ha encontrado una nueva clase de proteínas VQ, a las que hemos denominado proteínas R-VQ (*R protein-VQ*). Las proteínas codificadas por los genes *FvVQ13*, *FaVQ13C* y *FaVQ13D* portan dominios NB-ARC y LRR_8 característicos de proteínas R, asociadas comúnmente con la inmunidad innata y proteínas de resistencia a enfermedades en plantas. Aunque la existencia de este nuevo tipo de proteínas VQ ha podido ser verificada en otras especies, no es común en el reino vegetal. Ello sugiere unos mecanismos de aparición similares a los que originaron al grupo de las *R protein-WRKY*.
5. Los diferentes patrones de expresión y la expresión diferencial de los genes VQ en varios tejidos, así como en respuesta a las fitohormonas SA y MeJA y al hongo hemibiotrofo *Colletotrichum acutatum*, indican que cumplen diferentes funciones en la regulación de la fisiología y el desarrollo de la planta de fresa y, concretamente, en la maduración del fruto y los mecanismos de defensa.

6. El gen *FaWRKY1* es un regulador negativo de la resistencia del fruto de fresa frente a *Colletotrichum acutatum*. Sin embargo, podría actuar como regulador positivo frente a otros hongos y patógenos bacterianos, como en otras especies. Este papel dual en los mecanismos de resistencia frente a distintos patógenos podría ser ejercido interviniendo sobre el equilibrio de las dos principales rutas de defensa en plantas: las mediadas por ácido salicílico (SA) y ácido jasmónico (JA). De esta forma, se ha propuesto un modelo para la interacción *Colletotrichum*-fruto en el que *FaWRKY1* actuaría potenciando a la ruta de defensa mediada por SA. La regulación negativa sobre la ruta de JA, ejercida por otros genes (JAZ, cuya expresión sería regulada por *FaWRKY1*), así como el antagonismo entre las dos principales rutas de defensa, provocaría la inhibición de la ruta de JA y un incremento de la susceptibilidad frente a hongos necrotrofos.